Japanese Kokai Patent Application No. P2002-112149A

Job No.: 228-123905 Ref.: ORDER NO. 9159
Translated from Japanese by the McElroy Translation Company
800-531-9977 customerservice@mcelroytranslation.com

(19) JAPANESE PATENT OFFICE (JP)	(12) KOKAI TOKUHYO PATENT GAZETTE (A)	(11) PATENT APPLICATION PUBLICATION NO. 2002-112149

(43) Publication Date: April 12, 2002

(51) Int. Cl.[7]:	Identification Codes:	FI	Theme Codes (for reference)
H 04 N 5/76		H 04 N 5/76	A 5C025
5/445		5/445	Z 5C052
5/78		5/78	B 5C053
5/765		7/173	610 B 5C059
5/781		5/781	510 C 5C063

Examination Request: Not filed No. of Claims: 40 (Total of 36 pages; OL)

Continued on last page

(21) Filing No.: Japanese Patent Application No. 2000-301393

(22) Filing Date: September 29, 2000

(71) Applicant: 000002185
Sony Corp.
6-7-35 Kitashinagawa,
Shinagawa-ku, Tokyo

(72) Inventors: Shinji Negishi
Sony Corp.
6-7-35 Kitashinagawa,
Shinagawa-ku, Tokyo

(72) Inventor: Hideki Koyaragi
Sony Corp.
6-7-35 Kitashinagawa,
Shinagawa-ku, Tokyo

(74) Agents: 100067736
Akira Koike, patent attorney, and
2 others

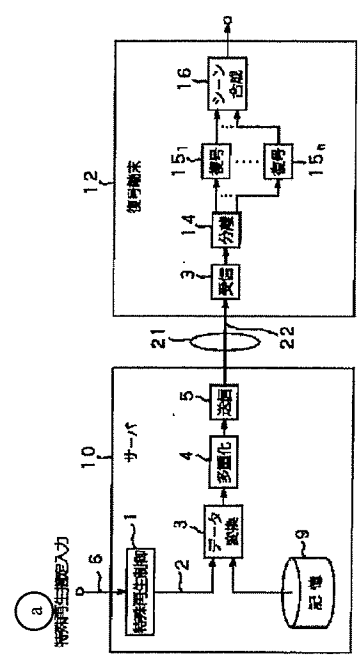
Continued on last page

(54) [Title] DATA PROCESSING METHOD AND DEVICE, DATA TRANSMISSION SYSTEM, AND TRANSMISSION MEDIUM

(57) Abstract

Problem: To provide a scheme characterized by the fact that when special playback is carried out on the decoding terminal side, decoding and display, etc., of data other than video data are possible, and scene description data can be delivered and decoded, and data can be delivered while the synchronization relationship between data is maintained, and the transmission bit rate or another evaluation standard is met.

Constitution: Data transformation part (7) of server (10) works as follows: when normal playback is carried out at decoding terminal (12), the data for use in said normal playback are output; when special playback is carried out at decoding terminal (12), time information pertaining to playback, such as the display start time, the display time, the display end time, etc., of AU of the data for use in normal playback is rewritten corresponding to the special playback, and it is decoded into data for output.



Key: a Special playback assignment input
1 Special playback control
3 Data transformation
4 Multiplexing
5 Sending
9 Storage
10 Server
12 Decoding terminal
3[sic; 13] Receiving
14 Separation
$15_1 \ldots 15_n$ Decoding
16 Scene synthesis

Claims

1. A data processing method characterized by the following facts: in a data processing method adopted for transmission of data encoded into prescribed coding units to the receiving side,

when normal playback is carried out on said receiving side, the data for use in said normal playback are output,

and, when special playback is carried out on said receiving side, time information pertaining to playback of said coding units of the data for use in said normal playback is transformed and output corresponding to said special playback.

2. The data processing method described in Claim 1, characterized by the following facts: when special playback is carried out on said receiving side, the coding units for output of the data used in said normal playback are selected such that an evaluation standard pertaining to restriction on the data that can be output within a prescribed time is met,

and the time information pertaining to playback of said selected coding units is transformed and output corresponding to said special playback.

3. The data processing method described in Claim 1, characterized by the fact that corresponding to said special playback, the coding units for transformation of said time information are transformed and output such that an evaluation standard pertaining to restriction on the data that can be output within a prescribed time is met.

4. The data processing method described in Claim 1, characterized by the fact that the time information pertaining to playback of said coding units includes the playback start time, playback time, playback end time, decoding time, and/or data arrival time when said coding units are played-back on said receiving side.

5. The data processing method described in Claim 1, characterized by the fact that corresponding to the playback speed in said special playback, the time information pertaining to playback of said coding units of the data for use in said normal playback is transformed.

6. The data processing method described in Claim 1, characterized by the fact that when jump playback is carried out as said special playback in which the playback site is moved to time-discontinuous coding units, output is stopped for the coding units corresponding to the start time and end time of said jump playback.

7. The data processing method described in Claim 1, characterized by the fact that said time information is changed so that the playback end time of coding units played-back extending across the start time of said special playback becomes the start time of said special playback.

8. The data processing method described in Claim 1, characterized by the fact that said time information is changed so that the playback start time of coding units played-back extending across the end time of said special playback becomes the end time of said special playback.

9. The data processing method described in Claim 6, characterized by the fact that when jump playback is carried out as said special playback in which the playback site to time-discontinuous coding units is moved, output is stopped for coding units played-back extending across the start time or end time of said jump playback.

10. The data processing method described in Claim 2, characterized by the fact that when said coding units are selected, a coding unit encoded for said data without using prediction between coding units is selected with priority.

11. The data processing method described in Claim 2, characterized by the fact that said evaluation standard is the bit rate of the output data.

12. The data processing method described in Claim 2, characterized by the fact that said evaluation standard is the degree of difficulty of decoding of scene description data when a scene is formed from said data.

13. The data processing method described in Claim 2, characterized by the fact that said evaluation standard is the degree of difficulty of scene synthesis and scene playback of scene description data when a scene is formed from said data.

14. A data processing device characterized by the fact that in a data processing device adopted for transmission of data encoded into prescribed coding units to the receiving side, there is a data transformation means that works as follows:

when normal playback is carried out on said receiving side, the data for use in said normal playback are output, and, when special playback is carried out on said receiving side, time information pertaining to playback of said coding units of the data for use in said normal playback is transformed and output corresponding to said special playback.

15. The data processing device described in Claim 14, characterized by the fact that said data transformation means works as follows: when special playback is carried out on said receiving side, the coding units for output of the data used in said normal playback are selected

such that an evaluation standard pertaining to restriction on the data that can be output within a prescribed time is met, and time information pertaining to playback of said selected coding units is transformed and output corresponding to said special playback.

16. The data processing device described in Claim 14, characterized by the fact that corresponding to said special playback, said data transformation means transforms and outputs coding units for transformation of said time information such that an evaluation standard pertaining to restriction on the data that can be output within a prescribed time is met.

17. The data processing device described in Claim 14, characterized by the fact that the time information pertaining to playback of said coding units includes the playback start time, playback time, playback end time, decoding time, and/or data arrival time when said coding units are played-back on said receiving side.

18. The data processing device described in Claim 14, characterized by the fact that corresponding to the playback speed in said special playback, said data transformation means transforms the time information pertaining to playback of said coding units of the data for use in said normal playback.

19. The data processing device described in Claim 14, characterized by the fact that when jump playback is carried out as said special playback in which the playback site to the time-discontinuous coding units is moved, output is stopped for the coding units corresponding to the start time and end time of said jump playback.

20. The data processing device described in Claim 14, characterized by the fact that said data transformation means changes the time information so that the playback end time of the coding units played-back extending across the start time of said special playback becomes the start time of said special playback.

21. The data processing device described in Claim 14, characterized by the fact that said data transformation means changes said time information so that the playback start time of coding units played-back extending across the end time of said special playback becomes the end time of said special playback.

22. The data processing device described in Claim 19, characterized by the fact that said data transformation means works as follows: when jump playback is carried out as said special playback in which the playback site is moved to time-discontinuous coding units, output is stopped for the coding units played-back extending across the start time or end time of said jump playback.

23. The data processing device described in Claim 15, characterized by the fact that when said coding units are selected, said data transformation means selects as priority a coding unit encoded for said data without using prediction between coding units.

24. The data processing device described in Claim 15, characterized by the fact that said evaluation standard is the bit rate of the output data.

25. The data processing device described in Claim 15, characterized by the fact that said evaluation standard is the degree of difficulty of decoding of scene description data when a scene is formed from said data.

26. The data processing device described in Claim 15, characterized by the fact that said evaluation standard is the degree of difficulty of scene synthesis and scene playback of scene description data when a scene is formed from said data.

27. A data transmission system characterized by the fact that in a data transmission system consisting of a transmitter that transmits data encoded into prescribed coding units and a receiver that receives said data, there is a data transformation means that works as follows: when normal playback is carried out on said receiving side, the data for use in said normal playback are output, and, when special playback is carried out on said receiving side, time information pertaining to playback of said coding units of the data for use in said normal playback is transformed and output corresponding to said special playback.

28. The data transmission system described in Claim 27, characterized by the fact that said data transformation means works as follows: when special playback is carried out on said receiving side, the coding units for output of the data used in said normal playback are selected such that an evaluation standard pertaining to the restriction on the data that can be output within a prescribed time is met, and the time information pertaining to playback of said selected coding units is transformed and output corresponding to said special playback.

29. The data transmission system described in Claim 27, characterized by the fact that corresponding to said special playback, said data transformation means transforms and outputs the coding units for transformation of said time information such that an evaluation standard pertaining to restriction on the data that can be output within a prescribed time is met.

30. The data transmission system described in Claim 27, characterized by the fact that the time information pertaining to playback of said coding units includes the playback start time, playback time, playback end time, decoding time, and/or data arrival time when said coding units are played-back on said receiving side.

31. The data transmission system described in Claim 27, characterized by the fact that corresponding to the playback speed in said special playback, said data transformation means transforms the time information pertaining to playback of said coding units of the data for use in said normal playback.

32. The data transmission system described in Claim 27, characterized by the fact that when jump playback is carried out as said special playback in which the playback site is moved

to time-discontinuous coding units, output is stopped for the coding units corresponding to the start time and end time of said jump playback.

33. The data transmission system described in Claim 27, characterized by the fact that said data transformation means changes the time information so that the playback end time of coding units played-back extending across the start time of said special playback becomes the start time of said special playback.

34. The data transmission system described in Claim 27, characterized by the fact that said data transformation means changes said time information so that the playback start time of coding units played-back extending across the end time of said special playback becomes the end time of said special playback.

35. The data transmission system described in Claim 32, characterized by the fact that said data transformation means works as follows: when jump playback is carried out as said special playback in which the playback site to the time-discontinuous coding units is moved, output is stopped for coding units played-back extending across the start time or end time of said jump playback.

36. The data transmission system described in Claim 28, characterized by the fact that when said coding units are selected, said data transformation means selects as priority a coding unit encoded for said data without using prediction between coding units.

37. The data transmission system described in Claim 28, characterized by the fact that said evaluation standard is the bit rate of the output data.

38. The data transmission system described in Claim 28, characterized by the fact that said evaluation standard is the degree of difficulty of decoding of scene description data when a scene is formed from said data.

39. The data transmission system described in Claim 28, characterized by the fact that said evaluation standard is the degree of difficulty of scene synthesis and scene playback of scene description data when a scene is formed from said data.

40. A transmission medium characterized by the following facts: the transmission medium is for transmission of data encoded into prescribed coding units at a transmitter to a receiver; for this transmission medium,

when normal playback is carried out at said receiver, data for use in said normal playback are transmitted,

when normal playback is carried out at said receiver, data for use in said normal playback are transmitted [sic], and, when special playback is carried out at said receiver, data obtained by transforming the time information pertaining to playback of said coding units of the data for use in said normal playback corresponding to said special playback are transmitted.

Detailed explanation of the invention

[0001]

Technical field of the invention

The present invention pertains to a data processing method, a data processing device, a data transmission system and a transmission medium that can be adopted as an optimum scheme in a data delivery system that carries out the following operation: multimedia data composed of still pictures, moving pictures, and other video data, audio data, text data, graphics data, etc., together with scene description data for forming a scene using said multimedia data are delivered; the delivered multimedia data and scene description data are received at a decoding terminal, and the data are decoded and displayed at the decoding terminal.

[0002]

Prior art

Figure 39 is a diagram illustrating an example of constitution of a conventional data delivery system that works as follows: video data, etc., prepared by compressing and storing the video signals of still pictures and moving pictures are delivered via a transmission medium, and they are received, decoded and displayed at the decoding terminal. In Figure 39, in order to simplify the explanation, only the path of the video data is shown. In the following, as an example, a case in which video data are delivered as packets of a transport stream (hereinafter to be referred to as TS) defined in ISO (International Organization for Standardization)/IEC (International Electrotechnical Commission 13818-1) (so-called MPEG2 Systems) will be explained.

[0003]

As shown in Figure 39, server (200) has storage part (209) that stores video data. Video data read from said storage part (209) are transformed into packets of a TS by means of multiplexer (204), and they are taken as delivery data (211) by means of transmitter (205) for output to transmission medium (210). For example, they are delivered to decoding terminal (212). In this case, transmission of said delivery data (211) of the TS is carried out by means of the protocol adopted in transmission medium (210). For example, a TS that meets the code of ISO/IEC 13818-1 can be transmitted according to the method defined in IEC 61883 "Digital interface for consumer audio/video equipment" using a transmission medium according to the code of IEEE (Institute of Electrical and Electronics Engineers) 1394. Also, one may adopt a scheme in which multiplexer (204) and transmitter (205) are an integrated unit.

[0004]

At said decoding terminal (212), said delivery data (211) are received by receiver (213) and are sent to separator (214). In separator (214), the video data are separated from said packets of the TS, and are sent to decoder (215). At decoder (215), the encoded video data are decoded. The decoded video data are sent to, e.g., a display unit not shown in the figure, and are displayed as video images.

[0005]

For said data delivery system, for example, when special playback/display, such as FFWD (fast-forward) play, frame-by-frame play, pause, etc., is carried out, for example, as the user operates the terminal front panel, remote controller, or the like, special playback assigning signal (206) (signal assigning FFWD play, frame-by-frame play, or the like) is input to special playback controller (216) of said decoding terminal (212). In this case, special playback controller (216) of decoding terminal (212) generates special playback request signal (220) for requesting the video data for special playback of the type assigned corresponding to special playback assigning signal (206) for server (200), and said special playback request signal (220) is transmitted via transmission medium (210) to special playback controller (201) of server (200).

[0006]

Said special playback controller (201) of server (200) that has received said special playback request signal (220) generates control signals (202a), (202b) upon said request, and the generated control signals are sent to corresponding multiplexer (204) and transmitter (205), respectively. Under control of special playback controller (201) by means of control signal (202b), multiplexer (204) reads from storage part (209) the video data for special playback that enable the type of special playback assigned by said user at decoding terminal (212). Here, multiplexer (204) transforms the video data for special playback into TS packets that are sent to transmitter (205). Under control of special playback controller (201) using control signal (202a), the packets of the video data for special playback are delivered as delivery data (211) to decoding terminal (212) by transmitter (205).

[0007]

At decoding terminal (212), when delivery data (211) composed of the video data for special playback are fed, said control signals (217a), (217b) for special playback control corresponding to said special playback assigning signal (206) are output from said special playback controller (216), and are sent to corresponding receiver (213) and decoder (215),

respectively. Under control of said special playback controller (216) and by means of control signal (217b), receiver (213) receives delivery data (211) composed of said video data for special playback, and sends them to separator (214). At said separator (214), the video data for special playback are separated from said packets of the TS, and are sent to decoder (215). At decoder (215), under control of special playback controller (216) and using control signal (217a), the video data for special playback are decoded. As a result, special playback display, such as FFWD play, frame-by-frame play, or the like, is carried out on a display unit not shown in the figure.

[0008]

In the decoding method of video frames defined by ISO/IEC 13818-2, there are the following pictures: I-pictures (Intra-coded pictures) encoded from only the data in a frame, B-pictures (Bidirectionally predictive-coded picture) encoded using intra-frame prediction, and P-pictures (Predictive-coded pictures). In the data delivery system shown in Figure 39, as the video data for said special playback read from storage part (209), I-pictures that do not use said intra-video-frame predictive processing are used. That is, in order to randomly access the video data for normal playback, I-pictures are contained on a regular basis, and I-pictures are extracted to form the video data for special playback. As explained above, in the conventional data delivery system shown in Figure 39, for example, when FFWD play or other special playback is carried out at decoding terminal (212), only the video data for special playback, such as video data made of I-pictures of ISO/IEC 13818-2, are delivered from server (200) to decoding terminal (212).

[0009]

On the other hand, in said data delivery system, when compressed video data according to ISO/IEC 13818-2 (so-called MPEG2 video) are delivered, the compressed video data defined in said ISO/IEC 13818-2 have to be encoded so that the decoder buffer does not undergo overflow or underflow. Also, the decoder buffer corresponds to an input buffer, not shown in the figure, in decoder (215). When more data than the size of the buffer defined in ISO/IEC 13818-2 are input, said decoder buffer overflows. On the other hand, when the data needed for decoding does not arrive at the time for decoding, a state of underflow exists.

[0010]

However, for video data comprising only I-pictures, such as said video data for special playback, the data quantity is large, and the decoder buffer may undergo overflow or underflow. Consequently, in the data delivery system of the prior art, special data for special playback that are different from those for normal playback should be prepared in advance, and, when special

playback is carried out at the decoding terminal, the special data for special playback should be delivered, so that overflow or underflow does not take place for the decoder buffer. Also, at the decoding terminal, a particular terminal should be arranged so that a particular special playback processing different from conventional special playback processing can be carried out corresponding to the particular data for special playback.

[0011]

That is, for a conventional data delivery system, in order to realize special playback free of overflow or underflow of the decoder buffer, particular data for special playback that are different from video data for special playback comprising only said I-pictures must be prepared in advance, and the particular data must be delivered in the special playback. Similarly, at the decoding terminal, a terminal having various particular decoders (215) that can correspond to particular data for special playback is required. Also, in special playback controller (216), processing of the data for special playback must be controlled for receiver (213), separator (214) and decoder (215).

[0012]

On this background, the present patent applicant proposed the following technology in Japanese Patent Application No. 2000-178999 and Japanese Patent Application No. 2000-179000: in a server, the video data for normal playback read from a storage part are used to carry out the type of special playback assigned by the user; the obtained data are transformed into video data that meet the code of ISO/IEC 13818-2, and the transformed video data are delivered to a decoding terminal; as a result, the use and advance preparation of the aforementioned particular delivery data for special playback are not required, and a particular decoding terminal that can cope with the particular delivery data for special playback is not required, so the constitution is simplified in this technology.

[0013]

Figure 40 is a schematic diagram illustrating the constitution of a data delivery system that works as follows: data obtained by carrying out special playback using the video data for normal playback are transformed into video data that meet the code of ISO/IEC 13818-2 for output. In the example shown in Figure 40, a case in which video data or the like are transformed into packets for delivery in the transport stream (TS) defined in ISO (International Organization for Standardization)/IEC (International Electrotechnical Commission) 13818-1 will be explained.

[0014]

As shown in Figure 40, server (220) has storage part (229) for storing multimedia data, such as video data of still pictures, moving pictures, etc., audio data, text data, graphics data, etc. From said storage part (229), for example, the video data are read, and the video data are sent via data transformation part (223) to be explained later to multiplexer (224). In multiplexer (224), the data output from data transformation part (223) are transformed into packets for a TS. Said TS packets are taken as delivery data (231) by transmitter (225) and are output to transmission medium (230), and, they are delivered to, e.g., decoding terminal (232). In this case, delivery data (231) of said TS are transmitted using the protocol adopted in transmission medium (230). For example, a TS that meets the code of ISO/IEC 13818-1 may be transmitted using the method defined in IEC 61883 "Digital interface for consumer audio/video equipment" and using a transmission medium defined according to, e.g., IEEE (Institute of Electrical and Electronics Engineers) 1394.

[0015]

At decoding terminal (232), said delivery data (231) are received by receiver (233), and they are sent to separator (234). In said separator (234), the video data are separated from said TS packets, and are then sent to decoder (235). In decoder (235), the fed data are decoded, that is, the encoded video data are decoded. The decoded video data are sent to a display unit not shown in the figure, and are displayed as video images.

[0016]

At decoding terminal (232) of said data delivery system, when a special playback display is carried out, for example, special playback assigning signal (226) corresponding to the operation carried out by the user of decoding terminal (232) is transmitted from a transmission medium interface part or the like not shown in the figure in decoding terminal (232) via transmission medium (230) to server (220). Said special playback assigning signal (226) is a signal that assigns the special playback type, such as FFWD play, REWIND play, frame-by-frame play, or the like, and the type of video data or the like stored in storage part (229). Also, if server (220) and decoding terminal (232) are connected to each other with a short distance between them using, e.g., a network for home use, and the user can work at the front panel or remote controller or the like of server (220), when the user works at the front panel or remote controller of said server (220), special playback assigning signal (226) also can be directly input to server (220).

[0017]

Said special playback assigning signal (226) input to said server (220) is input to special playback controller (221) arranged inside server (220). Corresponding to special playback assigning signal (226), said special playback controller (221) generates control signal (222) for special playback control containing assignment of the type of special playback and the video data, and sends it to data transformation part (223).

[0018]

Under control of special playback controller (221) and using control signal (222), data transformation part (223) reads video data from storage part (229). In addition, data transformation part (223) uses the video data read from storage part (229) to transform the data obtained as a result of the type of special playback assigned with said control signal (222) into video data that meet the code of ISO/IEC 13818-2. That is, when decoding is carried out just as in normal playback in decoder (235) of decoding terminal (232), said data transformation part (223) transforms the video data read from storage part (229) into video data for realizing the special playback, such as FFWD play, REWIND play, frame-by-frame play, etc. (special playback assigned by the user).

[0019]

In the following, a brief account will be presented on the data transformation processing in said data transformation part (223) with reference to Figures 41 and 42.

[0020]

Figure 41 is a schematic diagram illustrating data transformation processing when said data transformation part (223) transforms video data for normal playback encoded according to MPEG2 video (the video data read from storage part (229)) into video data that realize FFWD play as an example of special playback processing and that meet the code of ISO/IEC 13818-2. Here, in this figure, I represents the I-picture, P represents a P-picture, and B represents a B-picture. Here, according to a code of MPEG2 video, since inter-picture prediction is adopted in encoding, the coding sequence (the sequence of coding of data in the bit stream) and the actual display sequence may be different. Consequently, in Figure 41, both the coding sequence and the display sequence are shown. Here, Figure 41(a) shows the coding sequence of the video data for normal playback, and Figure 41(b) shows the display sequence when the video data for normal playback are decoded and displayed. In addition, Figure 41(c) shows the coding sequence in the case of transformation processing for special playback to go to FFWD play region FS after normal playback region US, followed by returning to normal playback region US. Figure 41(d)

shows the display sequence when transformation processing is carried out for the special playback shown in Figure 41(c).

[0021]

At data transformation part (223), for FFWD play region FS that carries out special playback, as indicated by E_k, E_m, and E_n in the figure, I-pictures (I_k, I_m, I_n) in the video data for normal playback shown in Figure 41(a) are extracted for use, and, in addition, in order to ensure that there is no problem in the decoder buffer, data transformation processing is carried out so that repeat picture B_R is inserted between said I-pictures. Here, said repeat picture B_R is a picture that repeats the prediction original picture, and this picture is handled as a B-picture in decoding. Here, insertion of repeat picture B_R has an effect of adjusting the speed of FFWD play.

[0022]

Similar to Figure 41, Figure 42 is a schematic diagram illustrating data transformation processing in the following operation: at said data transformation part (223), the video data for normal playback encoded according to MPEG2 video (the video data read from storage part (229)) are transformed into video data that realize REWIND play as an example of special playback processing and that meet the code of ISO/IEC 13818-2. Here, Figure 42(a) shows the coding sequence of the video data for normal playback, Figure 42(b) shows the display sequence when the video data for normal playback are decoded and displayed. Also, Figure 42(c) shows the coding sequence in the case of transformation processing for special playback to go to REWIND play region BS after normal playback region US, followed by returning to normal playback region US. Figure 42(d) shows the display sequence when transformation processing is carried out for the special playback shown in Figure 42(c).

[0023]

At data transformation part (223), for REWIND play region BS that carries out special playback, as indicated by E_k, E_m, and E_n in the figure, I-pictures (I_k, I_m, I_n) in the video data for the normal playback shown in Figure 42(a) are extracted and their order is reshuffled, and, in addition, in order to ensure that there is no problem in the decoder buffer, data transformation processing is carried out so that repeat picture B_R is inserted between said I-pictures.

[0024]

Here, the video data for special playback that are transformed by data transformation part (223) are delivered to decoding terminal (232) via the constitution of multiplexer (224) and thereafter as explained above.

[0025]

Problems to be solved by the invention

However, for television broadcasting in the prior art, one video signal is displayed on the screen of the image display unit, and only one audio signal is output from the speaker. However, in recent years, one scene may be composed of multimedia data, such as video data of still pictures, moving pictures, etc., audio data, text data, graphics data, etc. As a method for describing the constitution of a scene comprised of said multimedia data, there are the following schemes: HTML (HyperText Markup Language) for a homepage of the so-called internet, MPEG4 BIFS (Binary Format for the Scene) as a scene description scheme defined in ISO/IEC 14496-1, VRML (Virtual Reality Modeling Language) defined in ISO/IEC 14772, Java (trademark), etc. In the following, data describing a scene constitution will be called a scene description.

[0026]

In the following, an example of scene description using VRML and MPEG4 BIFS will be explained with reference to Figure 43. In Figure 43, the contents of the scene description are shown. According to VRML, the scene description is carried out by means of the text data shown in Figure 43, and, according to MPEG4 BIFS, the scene description is carried out by means of binary code transformed from the text data.

[0027]

The scene description of VRML and MPEG4 BIFS is represented by basic description units known as nodes. In the example shown in Figure 43, the nodes are represented by wide italic characters. Here, the nodes are units describing the displayed objects and the coupling relationship between the objects, and they contain data known as fields representing the node characteristics and attributes. For example, the Transform node shown in Figure 43 is a node that can assign three-dimensional coordinate transformation. By means of a translation field in the node, the translational movement distance of the origin of the coordinates is assigned. Also, other fields may be present that can assign a node in the field. For example, the transform node in Figure 43 is a children field indicating a children node group subjected to coordinate transformation by means of the transform node, and, by means of the children field, for example, shape nodes, etc., can be grouped. When the displayed objects are arranged on the scene, the nodes representing the objects are grouped together with nodes representing attributes. In addition, said nodes are grouped by means of nodes indicating the arrangement position. For

example, an object represented by the shape node in Figure 43 is arranged in the scene by adopting the translational movement assigned by the transform node as the parent node.

[0028]

Said video data and audio data, etc., are arranged for display in space and time by means of said scene description. For example, the Movie Texture node in Figure 43 specifies that the moving picture assigned by the ID known as 3 is bonded on the surface of a cube.

[0029]

Problems to be solved by the invention

As explained above, a scheme has been available in recent years wherein one scene is composed of multimedia data, such as video data of still pictures, moving pictures, etc., audio data, text data, graphics data, etc. However, in the data delivery system of the prior art, only the video data are decoded and displayed in special playback.

[0030]

Consequently, for example, even if multimedia data, such as video data of still pictures, moving pictures, etc., audio data, text data, graphics data, etc., are delivered, only the video data can be decoded and displayed in special playback. For example, even if data including data other than video data, such as audio data and text data for a subtitle, etc., are delivered, decoding and display for data other than the video data in special playback are still impossible in the data delivery system of the prior art.

[0031]

As a result, a desire exists for development of a scheme that can enable decoding and display of data other than video data, such as audio data, text data for a subtitle, etc., in special playback, such as FFWD play and REWIND play, etc.

[0032]

In the prior art, there is no scheme or means that can ensure that scene description data for forming said scene are also delivered and decoded during special playback. Consequently, for the data delivery system in the prior art, for example, even if said multimedia data are used to form a scene and the multimedia data are delivered, the scene still cannot be formed during special playback. As a result, the scenes displayed at the start and end of special playback are discontinuous, and this is undesirable.

[0033]

As explained above, a desire exists for development of a method and means for enabling delivery and decoding of said scene description data during special playback.

[0034]

In addition to realization of delivery, decoding, display, etc., of said multimedia data and scene description data during special playback, a synchronization relationship must be held among said data for display. In addition, data that meet the evaluation standard (the standard for ensuring no problem in the decoder buffer), such as the transmission bit rate, etc., must be delivered.

[0035]

The objective of the present invention is to solve the aforementioned problems of the prior art by providing a data processing method, a data processing device, a data transmission system, and a transmission medium characterized by the fact that when special playback is carried out, decoding, display, etc., of data other than the video data are possible, delivery, decoding, etc., of the scene description data are possible, and a synchronization relationship among data can be held while delivering the data that meet the transmission bit rate or another evaluation standard.

[0036]

Means to solve the problems

In order to solve the aforementioned problems, the present invention provides a data processing method characterized by the following facts: in a data processing method adopted for transmission of data encoded into prescribed coding units to the receiving side, when normal playback is carried out on said receiving side, the data for use in said normal playback are output, and, when special playback is carried out on said receiving side, time information pertaining to playback of said coding units of the data for use in said normal playback is transformed and output corresponding to said special playback.

[0037]

Also, the present invention provides a data processing device characterized by the fact that in a data processing device adopted for transmission of data encoded into prescribed coding units to the receiving side, there is a data transformation means that works as follows: when normal playback is carried out on said receiving side, the data for use in said normal playback are output, and, when special playback is carried out on said receiving side, time information

pertaining to playback of said coding units of the data for use in said normal playback is transformed and output corresponding to said special playback.

[0038]

Also, the present invention provides a data transmission system characterized by the fact that in a data transmission system consisting of a transmitter that transmits data encoded into prescribed coding units and a receiver that receives said data, there is a data transformation means that works as follows: when normal playback is carried out on said receiving side, the data for use in said normal playback are output, and, when special playback is carried out on said receiving side, time information pertaining to playback of said coding units of the data for use in said normal playback is transformed and output corresponding to said special playback.

[0039]

In addition, the present invention provides a transmission medium characterized by the following facts: the transmission medium is for transmission of data encoded into prescribed coding units at a transmitter to a receiver; for this transmission medium, when normal playback is carried out at said receiver, data for use in said normal playback are transmitted, when normal playback is carried out at said receiver, data for use in said normal playback are transmitted [sic], and, when special playback is carried out at said receiver, data obtained by transforming the time information pertaining to playback of said coding units of the data for use in said normal playback corresponding to said special playback are transmitted.

[0040]

That is, according to the present invention, it is possible to store and display the synchronization relationship between data even during special playback at the decoding terminal by converting the display time and the display period or display ending time of the display unit of the data for normal playback into data for special playback that are calculated and rewritten corresponding to the special playback. In this way, said problem can be solved. As a result, it is possible to deliver data that meet the bit rate or another evaluation standard even during special playback. In addition, according to the present invention, by transforming and outputting the display units in the data for normal playback so that the bit rate or another evaluation standard is met, it is possible to deliver data that meet the bit rate or another evaluation standard even during special playback.

[0041]

Embodiment of the invention

In the following, preferable embodiments of the present invention will be explained with reference to figures.

[0042]

Figure 1 is a diagram illustrating an example of the constitution of the data delivery system in an embodiment of the present invention in which multimedia data, such as video data of still pictures, moving pictures, etc., audio data, text data, graphics data, etc., as well as scene description data, etc., are delivered and are received and decoded for display at a decoding terminal. In the following explanation, as an example, video data, etc., are transformed into packets for delivery in the transport stream (TS) defined in ISO (International Organization for Standardization)/IEC (International Electrotechnical Commission) 13818-1 (the so-called MPEG2 system).

[0043]

As shown in Figure 1, server (10) has storage part (9) that stores multimedia data, such as video data of still pictures, moving pictures, etc., audio data, text data, graphics data, etc., and scene description data, etc. The data read from said storage part (9) are sent via, e.g., data transformation part (7) to be explained later to multiplexer (4). Said multiplexer (4) converts the data output from data transformation part (7) into TS packets. Said TS packets are then converted by transmitter (5) into delivery data (22) that are output to transmission medium (21) and distributed to decoding terminal (12). In this case, said delivery data (22) of the TS are transmitted according to the protocol adopted in transmission medium (21). For example, a TS that meets the code of ISO/IEC 13818-1 can be transmitted using the transmission medium defined in IEEE (Institute of Electrical and Electronics Engineers) 1394 according to the method defined in IEC 61883 "Digital interface for consumer audio/video equipment". Here, multiplexer (4) and transmitter (5) may be formed as an integrated unit.

[0044]

At decoding terminal (12), said delivery data (22) are received by means of receiver (13), and are sent to separator (14). At separator (14), the data are separated from said TS packets, and they are sent to the corresponding decoding parts among plural decoders (15_1)-(15_n), respectively. At said decoders (15_1)-(15_n), the fed data are decoded, that is, the encoded data are decoded here.

[0045]

When scene description data that describe the constitution of the scene are delivered, in scene synthesizer (16), the data decoded by means of said decoders (15_1)-(15_n) are synthesized according to said scene description data. The synthesized data prepared by means of scene synthesizer (16) are sent to a display unit or speaker or the like not shown in the figure, so that the scene image and sound are displayed/played. Here, plural decoding terminals (12) may be connected as well.

[0046]

In said decoding terminal (12) of the data delivery system, when a special playback display is carried out, for example, special playback assigning signal (6) corresponding to the operation carried out by the user on decoding terminal (12) is transmitted from a transmission medium interface part or the like not shown in the figure in decoding terminal (12) via transmission medium (21) to server (10). For example, said special playback assigning signal (6) contains an assignment of the type of special playback, such as FFWD play, REWIND play, frame-by-frame play, slow play, or the like as well as an assignment of the data stored in storage part (9). Also, if server (10) and decoding terminal (12) are connected together with a short distance between them using, e.g., a network for home use, and the user can operate the front panel or remote controller or the like of server (10), when the user operates the front panel or remote controller of said server (10), special playback assigning signal (6) also can be directly input to server (10).

[0047]

Said special playback assigning signal (6) input to said server (10) is input to special playback controller (1) arranged inside server (10). Corresponding to special playback assigning signal (6), said special playback controller (1) generates control signal (2) for special playback control containing assignment of the type of special playback and assignment of the data, and sends it to data transformation part (7). Here, there may be any number of said data transformation parts (7) corresponding to the number of delivered data.

[0048]

Under control of special playback controller (1) and using control signal (2), data transformation part (7) reads data from storage part (9). The data are then transformed into data for special playback to realize the type of special playback assigned by said control signal (2).

[0049]

In the following, the constitution and operation of data transformation part (7) in the data delivery system of the present embodiment will be explained in detail.

[0050]

Figure 2 is a diagram illustrating the detailed constitution of server (10) of the data delivery system having data transformation part (7) in Embodiment 1 of the present invention. Here, the operations of the various structural elements are the same as those above, except for the operation of data transformation part (7). Consequently, the aforementioned operations will not be explained again in detail.

[0051]

As shown in Figure 2, data transformation part (7) in Embodiment 1 has read part (17) for reading data from storage part (9) under control of control signal (2) from said special playback controller (1), and time information rewrite part (19) that rewrites encoded time information in the output data corresponding to the special playback. Here, when plural data transformation parts (7) are present, read parts (17) in data transformation parts (7) may have the same constitution for all of said data transformation parts (7).

[0052]

Said read part (17) reads from storage part (9) the data for normal playback assigned by control signal (2) from special playback controller (1) and sends same to time information rewrite part (19).

[0053]

Time information rewrite part (16) [sic; (19)] transforms the time information of the data for normal playback read by read part (17) from storage part (9) into time information of the data corresponding to said special playback, and encodes same in the output data. The time information of said data includes the data arrival time, display start time, display end time, display time, decoding time, etc. For the audio data, the aforementioned various types of time information pertain to the actual sound play. Here, in order to correlate the image display and sound play, said time information is adopted in representing the display start time, display end time, display time, etc. In said Embodiment 1, the data obtained by rewriting the time information by said time information rewrite part (16) are sent to said multiplexer (4).

[0054]

In the following, the time information transformation processing at time information rewrite part (19) of data transformation part (7) will be explained with reference to Figure 3. In the example shown in Figure 3, transformation processing is carried out for time information when FFWD play is carried out.

[0055]

Figure 3(a) is a diagram illustrating the display timing of data when the time information transformation processing for special playback by time information rewriting part (16) is not carried out for the data for normal playback read from storage part (9) (that is, when normal playback is carried out at decoding terminal (12)). In the coding method of a portion of MPEG2 video or the like, the actual display sequence and the coding sequence (the sequence of coding of data in the bit stream) may differ. In Figure 3, in order to facilitate understanding of the example, the display sequence is also shown. In Figure 3, AU30, AU31, AU32, etc., each represent one display unit of the data. They correspond to pictures for the video data. Coding of the data is usually carried out for each display unit. The display unit, that is, the coding unit, will be referred to as AU (access unit) below. Each AU starts display from display start time Ts, and ends the display at display end time Te after display time ΔT. Here, 1-AU display time Δ usually varies corresponding to the coding method.

[0056]

On the other hand, Figure 3(b) shows the display timing of the transformed data when special playback is carried out at decoding terminal (12), that is, when time information transformation processing for special playback (FFWD play in this example) is carried out by means of time information rewriting part (16) for the data for normal playback read from storage part (9). That is, Figure 3(b) shows the display timing when the FFWD play playback interval (special playback interval) starts in the middle of AU30' in the normal playback interval, AU31' represents the FFWD play interval and AU32' after said AU31' becomes the normal playback interval.

[0057]

Here, as an example of special playback, in Figure 3, FFWD play is carried out. In this case, the relationship between time point T in time t when the transformation processing for said special playback has not been carried out (hereinafter to be referred to as before-transformation time t) and time point T' in time t' when the transformation processing for the special playback

has been carried out (hereinafter to be referred to as after-transformation output time t') varies for different rounds of operation when said special playback is carried out.

[0058]

Consequently, at data transformation part (7) (time information rewrite part (19)) in this embodiment of the present invention, time point T' in said after-transformation output time t' can be computed using following equation (1) using special playback start time To' in said after-transformation output time t' and special playback start time To in said before-transformation time t (start time in said before-transformation time t corresponding to special playback start time To').

$$T' = To' + (T - To) / n \qquad (1)$$

Here, n in equation (1) represents the playback rate in the special playback. n has a value of 2 in the case of double-speed playback, and it has a negative value in REWIND play.

[0059]

On the other hand, in the normal playback, time point T' in said after-transformation output time t' can be computed using following equation (2) using special playback end time Ti' in said after-transformation output time t' and special playback end time Ti in said before-transformation time t (end time in said before-transformation time t corresponding to special playback start time [sic; end time] Ti').

$$T' = Ti' + (T - Ti) \qquad (2)$$

In the normal playback, there is no change in the preceding special playback end time. Consequently, the special playback start time at the start of the next special playback can be determined using equation (3) by means of equation (2).

$$To' = Ti' + (To - Ti) \qquad (3)$$

Using said equations (2)-(3), said data transformation part (7) can compute display start time Ts' of an AU and display end time Te' in after-transformation output time t' both during the normal playback and during the special playback based on display start time Ts and display end time Te of the AU in before-transformation time t. Also, display time ΔT' is computed either as 1/n times display time ΔT in before-transformation time t (n represents the playback rate), or it can be computed by subtracting display [start] time Ts' from display end time Te'.

[0060]

In this embodiment, data transformation part (7) is assigned with said special playback start time, special playback end time, and special playback rate n by means of said control signal

(2) and special playback controller (1). Also, said special playback start time, special playback end time, and special playback rate n may also be assigned by means of another data transformation part not shown in the figure. For example, the data delivery system in this embodiment may also have data transformation part (223) for transforming video data into a form for special playback as shown in said Japanese Patent Application No. 2000-178999 and Japanese Patent Application No. 2000-179000, Figure 40. When said data transformation part (223) is used to determine the special playback end time, the special playback start time, and the special playback speed matching the display timing of the video data, the special playback end time, the special playback start time and the special playback speed from data transformation part (223) may be directly assigned for data transformation part (7) in the present embodiment.

[0061]

For the data delivery system in the present embodiment, as explained above, in both normal playback and special playback, by computing said AU's display start time Ts' and display end time Te' in after-transformation output time t' as well as display time $\Delta T'$, it is possible to rewrite the display [start] time, display end time and display time encoded in the output data corresponding to the special playback. Also, when the decoding time, the data arrival time, and other time information are coded in the data, by means of time information rewrite part (19), said time information can be transformed into time information in after-transformation output time t' for output using equations (1) and (2).

[0062]

As explained above, according to the present embodiment, when special playback is executed at decoding terminal (12), the time information of the data for normal playback is transformed into time information of the data after transformation corresponding to said special playback, and the time information is encoded into data that are delivered from server (10). That is, in the data delivery system in the present embodiment, for the delivery data received by decoding terminal (12), because the time information for special playback has been transformed by means of server (10), special processing for special playback is not necessary at decoding terminal (12), and decoding and display, etc., of the result obtained by special playback can be obtained automatically. That is, in the present embodiment, for decoding terminal (12), no particular processing is required for special playback, and a particular terminal that can cope with particular delivery data for special playback is not required. In addition, according to the present embodiment, plural delivered data are transformed corresponding to the playback speed, so an offset in synchronization between the plural data is not present, and offset does not accumulate.

[0063]

In the following, processing for transforming the time information in time information rewrite part (19) when slow play is carried out as the special playback will be explained with reference to Figure 4 that is similar to Figure 3.

[0064]

Similar to Figure 3(a), Figure 4(a) is a diagram illustrating the display timing of data for normal playback at before-transformation time t. In Figure 4, AU40, AU41, AU42, etc., each represent one display unit of the data. Just like Figure 3(b), Figure 4(b) shows the display timing of the transformed data when special playback (slow play in this example) is carried out by means of time information rewriting part (16). That is, Figure 4(b) shows the display timing, in which the slow play interval starts in the middle of AU40' in a normal playback interval; AU41' represents a slow play interval, and AU42' succeeding said AU41' represents a normal playback interval.

[0065]

Here, for example, 0.5-times-speed playback is carried out as the special playback, that is, a value of 0.5 for playback rate n is used for computation in said equation (1) for data transformation part (7) (time information rewrite part (19)) in this embodiment of the present invention.

[0066]

In the example shown in Figure 4, if the playback speed is either the normal playback speed or a lower speed as the special playback, the time information transformation processing carried out at data transformation part (7) in this embodiment is as effective as mentioned previously. At decoding terminal (12), the same decoding and display can be carried out without any particular processing for slow playback, and a display result of slow playback can be obtained.

[0067]

In the following, with reference to Figure 5, which is similar to Figure 3, time information transformation processing at time information rewrite part (19) will be explained when special playback is carried out as a jump in the playback position to a display unit that is discontinuous in time.

[0068]

Just like Figure 3(a), Figure 5(a) represents the display timing of data for normal playback at before-transformation time t. In Figure 5, AU50, AU51, AU52, etc., each represent one display unit of data. On the other hand, like Figure 3(b), Figure 5(b) shows the display timing of the transformed data when special playback is carried out at said time information rewriting part (16) (jump in this example). That is, Figure 5(b) shows the display timing in a case in which a jump occurs in the middle of AU50' in a normal playback interval. Here, AU51 is not output during the period between special playback start time To' as the jump start time and special playback end time Ti' as the jump end time. After special playback start time To' in said AU50', AU51' after special playback end time Ti' is output.

[0069]

Here, in the case of a jump, a playback speed does not exist in this special playback. Consequently, special playback controller (1) assigns the special playback start time and the special playback end time for data transformation part (7). For said special playback start time, because conversion is possible for special playback start time To in before-transformation time t according to equation (3) and special playback start time To' in after-transformation output time t', it may be assigned at any time before or after said transformation. On the other hand, with regard to the special playback end time, special playback end times Ti and Ti' are assigned for both the time before transformation and the time after transformation. Here, if special playback start time To' in after-transformation output time t' is equal to special playback end time Ti', it is alright if Ti' is not assigned.

[0070]

In the example shown in Figure 5, data transformation part (7) does not output AU51 between special playback start time To' and special playback end time Ti' of the jump, and the time information of AU50 displayed extending across special playback start time To' may be changed and output so that the display end time becomes To', or it may not be output. In addition, the time information of AU52 displayed extending across special playback end time Ti' of the jump is changed and output so that the display time becomes Ti', or it may not be output.

[0071]

As can be seen from the example shown in Figure 5, even if said jump for moving the playback site to a display unit that is discontinuous in time is carried out as special playback, the transformation processing of the time information at data transformation part (7) in the present embodiment is as effective as mentioned previously, so the display result of a jump operation or

the like can be obtained by carrying out the same decoding and display or the like as in the normal playback without any particular processing for the jump at decoding terminal (12).

[0072]

Also, according to the present invention, by means of transformation corresponding to the special playback, scene description data that describe the constitution of a scene can be delivered and decoded, etc., even during special playback. Consequently, the problem of, e.g., a display of discontinuity of a scene at the start or end of special playback can be avoided.

[0073]

In said example, when the display time and decoding time and other time information are coded and added to the data, time information rewrite part (19) of data transformation part (7) rewrites the time information and outputs it. However, other schemes may also be adopted. For example, when time information is added to the data by multiplexer (4), change in the time information is notified from data transformation part (7) to multiplexer (4), and multiplexer (4) adds the changed time information to the data. Also, when time information is added to the data by means of transmitter (5), similarly, the change in the time information is notified from data transformation part (7) to transmitter (5), and transmitter (5) adds the changed time information. This can also be adopted for other embodiments to be explained later.

[0074]

Here, in the data delivery system where multimedia data, such as video data of still pictures, moving pictures, etc., audio data, text data, graphics data, etc., and scene description data are delivered, decoded and displayed, there is a requirement that the data be delivered while meeting the bit rate or another evaluation standard even in special playback.

[0075]

That is, as shown in the example illustrated in Figure 3, the delivery data in FFWD play are compressed on the time axis more than the delivery data in normal playback, and their average bit rate is higher than that of normal playback. On the other hand, for a system that delivers data via the transmission medium like that of the present embodiment, an upper limit of the bit rate allowed in delivery corresponding to the transmission capacity of the transmission medium and the decoding ability is determined. For example, if the bit rate of the delivery data is over the upper limit of the bit rate allowed in said delivery, data delay and loss will take place. In such case, by applying a restriction on the bit rate of said delivery data, the bit rate of the

delivery data can be prevented from exceeding the upper limit of the bit rate allowed in said delivery.

[0076]

For example, when the data contained in the delivery data within a certain time is relatively increased, the difficulty in decoding, scene synthesis and display increases, and correct display at the decoding terminal becomes difficult. In such case, by adding a restriction on the difficulty in decoding, scene synthesis and display of said delivery data, the difficulty can be decreased such that correct display at the decoding terminal is possible.

[0077]

In consideration of this problem, in Embodiment 2 of the present invention, data delivery can be carried out so that the bit rate or another evaluation standard is met even in the special playback. As a result, delay or loss of data can be prevented, and correct display of a scene at the decoding terminal is possible.

[0078]

Figure 6 is a diagram illustrating in detail server (10) of a data delivery system having data transformation part (7) in Embodiment 2 in the present invention.

[0079]

As shown in Figure 6, data transformation part (7) has read part (17) that reads the data from said storage part (9) based on control of control signal (2) from said special playback controller (1), time information rewrite part (19) that rewrites the time information coded in the output data corresponding to the special playback, as well as scheduler (18) that selects the AU for output based on the bit rate or another evaluation standard. Also, with data transformation part (7), transformation of the time information from the time of the data for normal playback before transformation to the time after transformation is carried out, and the time information is coded in the data for output. This processing is the same as that in Embodiment 1.

[0080]

In the following, transformation processing in scheduler (18) of data transformation part (7) in Embodiment 2 will be explained with reference to Figures 7 and 8.

[0081]

Figure 7 provides the same representation as that in Figure 3. Just like Figure 3(a), Figure 7(a) shows the display timing of normal playback data at before-transformation time t. In Figure 7, AU70, AU71, AU72, AU73, etc., each represent one display unit of the data. Just like Figure 3(b), Figures 7(b), (c), (d) show the display timing of the transformed data when an AU is selected corresponding to the bit rate allowed in delivery by means of scheduler (18) in this embodiment, while time information transformation processing is carried out for special playback (jump in this example) by means of time information rewrite part (16). That is, Figure 7(b) shows the display timing when AU71 and AU72 are selected by scheduler (18) in a FFWD play interval (special playback interval), and, at the same time, said AU71 and AU72 are subjected to time information transformation processing by means of time information rewrite part (16) to become AU71', AU72', and the succeeding AU73' is taken as a normal playback interval. Also, Figure 7(c) shows the display timing in a case in which only AU71 is selected by scheduler (18) in the FFWD play interval, and, at the same time, said AU71 is transformed by time information rewrite part (16) into AU71'. On the other hand, AU72 is not output. Then, the succeeding AU73' is taken as a normal playback interval. Figure 7(d) shows the display timing in a case in which AU71 or AU72 is not selected by scheduler (18) in a FFWD play interval, and the succeeding AU73' is taken as a normal playback interval.

[0082]

Here, as shown in Figure 7(a), in the special playback interval (FFWD play interval) at before-transformation time t, two AUs exist, that is, AU71 and AU72. In said Embodiment 1, the time information of AU71, AU72 is transformed corresponding to the special playback speed, and is output as AU71' and AU72'. However, as shown in Figure 8, when the special playback (FFWD play in the example shown in Figures 7 and 8) is carried out, the bit rate of the delivery data varies corresponding to the playback speed. When the bit rate varying in this way exceeds the tolerable bit rate of the transmission medium and the decoding terminal, data delay and loss, etc., take place.

[0083]

Here, scheduler (18) arranged in data transformation part (7) in this embodiment selects AU for output and AU not for output so that the bit rate allowed for the delivery data is not exceeded. For example, when the bit rate allowed for the delivery data is over bit rate B_R81 when only AU71 is output while AU72 is not output, and below bit rate B_R80 when both AU71 and AU[7]2 are output, scheduler (18) determines that AU72 is not output. The transformation output in this case is shown in Figure 7(c). Here, when the bit rate allowed for the delivery data is lower

than the bit rate B_R81 when only AU71 is output while AU72 is not output, scheduler (18) directs that neither AU71 nor AU72 be output. Figure 7(d) shows the transformation output in this case. On the other hand, when the bit rate allowed for the delivery data is over bit rate B_R80 when both AU71 and AU72 are output, scheduler (18) directs that both AU71 and AU72 be output. Figure 7(a) shows the transformation output in this case. Then, by means of time information rewrite part (19), for the AU selected and output by scheduler (18), as explained above, the time information is transformed based on the playback speed of the special playback.

[0084]

As explained above, according to Embodiment 2, by selecting and outputting display unit (AU) in the data for normal playback so that the bit rate or another evaluation standard is met, the delivery of data can meet the bit rate or another evaluation standard in the special playback. Here, the evaluation standard is not limited to the bit rate. For example, the evaluation standard may also be the number of polygons allowed in a certain time, the number of nodes in the scene description data, or another difficulty indicating data decoding, scene synthesis, display, etc. Also, the evaluation standard may be the number of characters in the text data and other data that can be output in a certain time.

[0085]

In addition, in Embodiment 2 of the present invention, data transformation part (7) may also work as follows: when a display unit (AU) for output and a display unit not for output are selected as explained above, display units encoded without using data in predicting between display units are output as priority, while the display units encoded by means of prediction are selected to not be output. As a result, at the decoding terminal, predictive decoding is possible with display units coded without using said prediction taken as the prediction elements.

[0086]

In said Embodiment 2, as an example, depending on whether an AU is selected for output, delivery data that meet the bit rate or other evaluation standard can be output. On the other hand, in Embodiment 3 to be presented below, the following scheme may be adopted: by means of transformation of the contents of AUs themselves, delivery data that meet the bit rate or another evaluation standard are output.

[0087]

Figure 9 shows in detail the constitution of server (10) of the data delivery system in Embodiment 3 of the present invention.

[0088]

As shown in Figure 9, server (10) has the same configuration as that in said Embodiments 1 and 2, except that it has filter (23) in the output section of data transformation part (7) corresponding to any of said embodiments.

[0089]

Said filter (23) transforms the data to a form for special playback by means of data transformation part (7) in said Embodiment 1 or 2, that is, the AU itself, to a form that meets the bit rate or another evaluation standard. Here, there may be a plurality of data transformation parts (7) and filters (23). That is, for filter (23) in said Embodiment 3, instead of only selection of AUs for output and AUs not for output as adopted by data transformation part (7) in Embodiment 2, the AU itself is transformed such that data that meet the bit rate or another evaluation standard are output. For example, for the text data, by decreasing the number of characters contained in each AU, the quantity of data for delivery can be decreased, and the data are transformed to meet the desired bit rate and are output.

[0090]

In the present embodiment, by means of transformation of an AU itself, delivery of data that meet the bit rate or another evaluation standard is possible even in special playback. Here, for an AU input to filter (23), because the time information has been transformed corresponding to the special playback by means of data transformation part (7) in Embodiment 1 or Embodiment 2, particular processing for special playback on the side of decoding terminal (12) is not necessary, and display for special playback can be realized automatically by carrying out the same decoding and display or the like just like that in normal playback free of any particular processing for the special playback at decoding terminal (12).

[0091]

In the following, examples of filter (23) will be explained.

[0092]

As the first example of said filter (23), for example, the data in a scene description are handled for individual divided units, and the scene description is transformed and output to each divided unit so that the transmission capacity or another evaluation standard is met. Filter (23) in the first example is used along with data transformation part (7) in Embodiment 1 or

Embodiment 2 of the present invention, and data that meet the bit rate or another evaluation standard can be delivered in the special playback.

[0093]

In the following, the operation of filter (23) in said first example of Embodiment 3 of the present invention will be explained.

[0094]

Said filter (23) of the first example transforms an input scene description based on hierarchical information. Said filter (23) obtains decoding terminal information indicating the decoding and display ability of decoding terminal (12) when the scene description is output. Said decoding terminal information shows the decoding and display ability of decoding terminal (12), such as the image frame when the scene description is displayed at decoding terminal (12), the upper limit of the node number, the upper limit of the number of polygons, the upper limit of the contained audio and video or other multimedia data, etc. Hierarchical information with information representing the transmission capacity of transmission medium (22) [sic; (21)] for delivery of the scene description in addition to the decoding terminal information is input to filter (23). Based on said hierarchical information, filter (23) transforms the scene description input into scene description data that form a hierarchical structure.

[0095]

For the data delivery system in Embodiment 3 having filter (23) of the first example, as explained above, by means of transformation of the scene description based on the hierarchical information, scene description data appropriate for transmission medium (22) for delivery can be delivered, and a scene description matching the performance of decoding terminal (12) can be delivered.

[0096]

In the following, the procedure of scene description transformation processing in filter (23) will be explained with reference to Figure 10.

[0097]

As shown in Figure 10, first, as step S200, filter (23) divides the scene description into division candidate units to be explained later. As shown in Figure 10, n represents the No. of a division candidate [unit]. Here, in order to transform the input scene description into scene description data made of plural hierarchical layers, a hierarchical layer of the output scene

description data is represented by m. Said hierarchical layer No. m starts at 0, and a smaller No. represents a more basic hierarchical layer.

[0098]

Then, in step S201, based on the hierarchical information, said filter (23) judges that division candidate n can be output as the current hierarchical layer. For example, when the byte number of the data allowed for the current hierarchical layer is restricted by the hierarchical information, whether the output scene description of the current hierarchical layer is smaller than the aforementioned limited byte number even when said division candidate n is added is checked. In step S201, when it is judged that division candidate n cannot be output to the current hierarchical layer, the flow goes to step S202. On the other hand, if output is possible, the flow goes to step S203.

[0099]

After going to step S202, filter (23) increases hierarchical layer No. m by one. That is, output to the current hierarchical layer No. m is ended, and then output is made to a new hierarchical layer of the scene description data. Then, the flow goes to step S203.

[0100]

After going to step S203, filter (23) outputs division candidate n to the current hierarchical layer No. m. Then, the flow goes to step S204.

[0101]

After going to step S204, filter (23) judges whether all of the division candidates have been processed. If YES, transformation processing is ended. On the other hand, if some division candidates still remain, the flow goes to step S205.

[0102]

When the flow goes to step S205, filter (23) increases division candidate n by one. That is, the next division candidate is taken as the processing object. Then the processing is repeated from step S201.

[0103]

In the following, with reference to Figure 11, MPEG4 BIFS will be taken as an example for explanation of division in the scene description transformation processing by filter (23) shown in Figure 10.

[0104]

First, the contents of the scene description data shown in Figure 11 will be explained. Then, division in the scene description processing in filter (23) will be explained.

[0105]

As shown in Figure 11, [T]ransform node (302) can assign a three-dimensional coordinate transformation, and the translational movement distance of the coordinate origin on its translation field (303) can be assigned. On the field, there is also a field that can assign other nodes, and the constitution of the scene description has the tree structure shown in Figure 12. As shown in Figure 12, the ellipse indicates the nodes, and the broken lines between nodes represent the propagation path of events, while the solid lines between nodes represent the parent-children relationship of the nodes. With respect to the parent node, the nodes representing the field of the parent node are called children nodes. For example, for Transform node (302) in Figure 11, there is children field (304) that shows the children node group that is coordinate-transformed by the Transform node, and Touch Sensor node (305) and Shape node (306) are grouped as children nodes. In this way, the node for grouping the children nodes in the children field is called a grouping node. A grouping node is defined in ISO/IEC 14772-1, Chapter 4.6.5, and it refers to a node having a field comprising a list of nodes. As defined in ISO/IEC 14772-1, Chapter 4.6.5, there is also a special exception that the field name may not be children. In the following explanation, such exception is also taken as included in the children field.

[0106]

When an object is to be set in a scene, the node representing the object is grouped together with a node representing attributes, and it is further grouped by means of a node indicating the arrangement position. The object representing Shape node (306) in Figure 11 is arranged in the scene by adopting a translational movement assigned by Transform node (302) as its parent node. The scene description shown in Figure 11 contains Sphere node (307) representing a ball, Box node (312) representing a cube, Cone node (317) representing a cone, and Cylinder node (322) representing a cylinder. Figure 13 shows the result of decoding and display of the scene description in this example.

[0107]

The scene description also may contain user interaction. ROUTE in Figure 11 represents propagation of an event. ROUTE (323) works as follows: when the touch time field of Touch Sensor node (305) with an identifier of 2 allotted to it varies, its value is taken as the event, and

propagation is made to the start time field of Time Sensor node (318) with 5 allotted as an identifier. In VRML, the identifier is represented by any character sequence following DEF as the keyword, and, in MPEG4 BIFS, a value known as node ID (node ID) is used as the identifier. When the user selects Shape node (306) grouped into children field (304) of Transform node (302) as the parent node, Touch Sensor node (305) outputs the selected time as a touch time event. A sensor that is grouped together with an annexed Shape node by means of said grouping node will be referred to as a Sensor node. The Sensor node in VRML is a Pointing device sensor defined in ISO/IEC 14772-1, Chapter 4.6.7.3, and the annexed Shape node refers to the Shape node grouped with the parent node of the Sensor node. On the other hand, Time Sensor node (318) outputs a period of 1 sec from the start time as a fraction_changed [e]vent.

[0018]

By means of ROUTE (324), a fraction_changed event representing the lapsed time output from Time Sensor node (318) is propagated to the set_fraction field of Color Interpolator node (319) with 6 as the identifier allotted to it. Said Color Interpolator node (319) has the function of linearity correction for the value of the RGB color space. For the key and key value field of Color Interpolator node (319), when the value of the input set_fraction field is 0, an RGB value of [000] is event-output as value_changed, and, when the value of the input set_fraction field is 1, RGB value [111] is event-output as value_changed. When the value of the input set_fraction field is between 0 and 1, the value after linearity correction between the RGB values of [000] and [111] is event-output as value_changed. That is, when the value of the input set_fraction field is 0.2, the RGB value [0.2 0.2 0.2] is event-output as value_changed.

[0109]

By means of ROUTE (325), the value value_changed of the linearity interpolation result is propagated to the diffuse color field of Material node (314) with 4 allotted to it as the identifier. Here, diffuse Color represents the diffusion color of the surface of the object represented by Shape node (311) where Material node (314) belongs. By means of event propagation via said ROUTE (323), ROUTE (324) and ROUTE (325), in the period of 1 sec from just after selection of a ball represented by the user, user interaction is realized with the RGB value of the displayed cube changed from [000] to [111]. This user interaction is represented by nodes correlated with propagation of the event indicated by said ROUTE (323), ROUTE (324), ROUTE (325), and the wide line frame in Figure 12. The data in the scene description needed in the user interaction is called the data needed for event propagation. Here, nodes other than those indicated by the wide line frame are not irrelevant to the event.

[0110]

For the scene description data shown in Figure 11 presented as an example above, they are divided by filter (23) of the first example in this embodiment to division candidate units of the scene description in step S200 shown in Figure 10.

[0111]

Here, because a so-called Node Insertion command is used, the children field of the grouping node is taken as the division unit. Here, when the data needed for event propagation for user interaction are not divided, there are three division candidates D0, D1, D2 shown in Figure 11.

[0112]

The division unit containing Group node (300) as the most-significant node in the input scene description is taken as the division candidate D0 with n = 0. The node below Transform node (315) is taken as division candidate D1 with n = 1. Shape node (316) in division candidate D1 with n = 1 is a children field of Transform node (315) as the grouping node, so it can be taken as another division candidate.

[0113]

In this example, Transform node (315) does not have a children field other than Shape node (316). Consequently, Shape node (316) is not taken as another division candidate. The node after Transform node (320) is taken as division candidate D2 with n = 2. Similarly, that following Shape node (321) may also be taken as another division candidate.

[0114]

Division candidate D0 with n = 0 is output to hierarchical layer m = 0 for sure. In step S201 shown in Figure 10, division candidate D1 with n = 1 judges whether output to the hierarchical layer with m = 0 is possible based on the hierarchical information.

[0115]

Figure 14 shows an example of judgment when the data quantity allowed for the hierarchical layer of the output scene description data is assigned according to the hierarchical information. In example A in Figure 14, when division candidate D1 with n = 1 is also output to hierarchical layer m = 0, the data quantity allowed for hierarchical layer m = 0 increases, so it is judged that division candidate D1 with n = 1 cannot be output to hierarchical layer m = 0.

[0116]

Consequently, with the procedure of step S202 shown in Figure 10, it is determined that the output of hierarchical layer m = 0 shown in B in Figure 14 only contains division candidate D0 with n = 0, and thereafter is output to hierarchical layer m = 1. With the procedure of step S203, division candidate D1 with n = 1 is output to hierarchical layer m = 1.

[0117]

When the same procedure of operation is carried out for division candidate D2 with n = 2, as indicated by A in Figure 14, even when division candidate D2 with n = 2 is output to hierarchical layer m = 1, the data quantity does not exceed the data quantity allowed for the sum of hierarchical layer m = 0 and hierarchical layer m = 1. Consequently, as indicated by C in Figure 14, it is determined that division candidate D2 with n = 2 is output to hierarchical layer m = 1, just like division candidate D1 with n = 1.

[0118]

With said procedure, filter (23) transforms the input scene description into scene description data composed of two hierarchical layers including the transformed scene description data output with hierarchical layer m = 0 shown in B in Figure 14 and the transformed scene description data output with hierarchical layer m = 1 shown in C in Figure 14.

[0119]

As an example of transformation of the scene description shown in A in Figure 15, as a result of transformation based on different hierarchical information with respect to the same input scene description as that in A of Figure 14, transformation is made to scene description data composed of three hierarchical layers.

[0120]

That is, just like the case shown in Figure 14, the scene description shown in A in Figure 15 is transformed into the transformed scene description data output with hierarchical layer m = 0 shown in B in Figure 15, the transformed scene description data with hierarchical layer m = 1 shown in C in Figure 15, and transformed data output with hierarchical layer m = 2 indicated by D in Figure 15.

[0121]

In this example of a transformation result, the transmission capacity of the transmission medium used in delivery of the scene description is low, and, for a transmission medium that can

transmit only a data quantity that allows hierarchical layer $m = 0$, only the scene description data of hierarchical layer $m = 0$ indicated in B in Figure 15 are transmitted.

[0122]

Even for a scene description with hierarchical layer $m = 0$, the data needed for event propagation for user interaction are not divided, so the same user interaction as that before transformation can be realized at decoding terminal (12).

[0123]

Also, for a transmission medium that has a sufficient transmission capacity with respect to the total data quantity of hierarchical layer $m = 0$ and hierarchical layer $m = 1$, the scene description data of both hierarchical layer $m = 0$ shown in B of Figure 15 and hierarchical layer $m = 1$ shown in C of Figure 15 are delivered.

[0124]

Because the scene description data with hierarchical layer $m = 1$ are inserted in the scene description with hierarchical layer $m = 0$ by means of a Node Insertion command, decoding and display of the scene description just like before transformation are possible at decoding terminal (12).

[0125]

In the first example, filter (23) can even cope with a case in which the transmission capacity of transmission medium (22) varies when a scene description is transformed based on hierarchical information that varies with time. Also, the same effect as that when the transformed scene description data are recorded in transmission medium (22) can be realized.

[0126]

In the example of the transformation result shown in Figure 15, the decoding and display ability of decoding terminal (12) that decodes and displays the scene description is low, and, for decoding terminal (12) that allows decoding and display of data only up to the data quantity allowed for hierarchical layer $m = 0$, only the scene description data with hierarchical layer $m = 0$ indicated by B in Figure 15 can be delivered. Even for a scene description with only hierarchical layer $m = 0$, the data needed for event propagation for user interaction are not divided, so the same user interaction as that before transformation can be realized at decoding terminal (12).

[0127]

For decoding terminal (12) that has sufficient ability of decoding and display with respect to the data quantity as the sum of the hierarchical layer m = 0 and hierarchical layer m = 1, the scene description data of the hierarchical layers with m = 0 as shown in B of Figure 15 and m = 1 shown in C of Figure 15 are distributed.

[0128]

Because the data of scene description (100) with hierarchical layer m = 1 are inserted into the scene with hierarchical layer m = 0 by means of a Node Insertion command, the same scene description as that before transformation can be decoded and displayed at decoding terminal (12).

[0129]

As explained above, by means of first filter (23), the scene description is transformed based on decoding terminal information that varies over time, so even if dynamic variation occurs in the decoding and display ability of decoding terminal (12) or if a decoding terminal (12) with new capability is added to the objects of delivery, the situation can be handled.

[0130]

In MPEG4 BIFS, in order to obtain a hierarchical scene description, a command for inserting a node may be used, or an Inline node may be used. Also, EXTERNPROTO described in ISO/IEC 14772-1, Chapter 4.9 may be used. Here, EXTERNPROTO refers to a method in which a node is defined according to the node definition method known as PROTO in the external scene description data, and the same EXTERNPROTO as that of VRML can also be used in MPEG4 BIFS.

[0131]

Also, the following scheme may be adopted: DEF/USE described in ISO/IEC 14772-1, Chapter 4.6.2 has the name DEF attached at a node, and the DEF-attached node can be used as a reference by means of USE from another site during a scene description.

[0132]

Also for MPEG4 BIFS, an identifier known as a node ID is attached to a node just like DEF, and, by assigning the node ID from another side during scene description, the reference adopted for USE can be used for VRML.

[0133]

Consequently, when a scene description is transformed to a hierarchical form, if the portion using DEF/USE described in ISO/IEC 14772-1, Chapter 4.6.2 is not divided into different division candidates, the reference relationship from USE to DEF nodes is not damaged, and scene description transformation is possible.

[0134]

In the examples shown in Figures 14 and 15, the data quantity allowed for each hierarchical layer is used as hierarchical information. However, any information that can judge whether the division candidate in the scene description can be contained in the scene description data of a certain hierarchical layer may be used. Examples include the upper limit of the node number contained in the hierarchical layer, the number of polygon data in computer graphics contained in the hierarchical layer, the limit on media data, such as audio data, video data, etc., contained in the hierarchical layer, or a combination of plural hierarchical information items.

[0135]

As explained above, for filter (23) in the first example, by transforming the input scene description into scene description data comprised of a plural hierarchical structure, the hierarchical structure of the scene description can be used for the purpose of conservation of transmission capacity when the scene description is transmitted.

[0136]

Also, according to filter (23) of the first example, by transforming the scene description into scene description data comprised of plural hierarchical layers and deleting only the scene description data of the hierarchical layers up to the data quantity to be deleted when data are deleted, a portion of the contents information that describes the scene description can be stored.

[0137]

In addition, the aforementioned scheme can be efficiently adopted in any scene description method involving division independent of the type of scene description method.

[0138]

In the following, the operation of filter (23) in said second example of Embodiment 3 of the present invention will be explained.

[0139]

As shown in Figure 16, filter (23) in said second example has scene description processor (24), ES (Elementary Stream) processor (25), and controller (26) that controls the operation of said parts. By means of scene description processor (24), the scene description data are changed, and, at the same time, multimedia data other than the scene description data can be changed by means of ES processor (25). Here, ES processor (25) is for carrying out transformation of the data into data with a different bit rate for re-coding corresponding to the transmission capacity and the ability of the decoding terminal. On the other hand, scene description processor (24) adjusts the data quantity by transforming the contents of the scene description corresponding to the transmission capacity of transmission medium (22) and the processability of decoding terminal (12). By using filter (23) having scene description processor (24) and ES processor (25) together with data transformation part (7) in Embodiment 1 or 2, data meeting the bit rate or another evaluation standard can be delivered even in special playback. Although not shown in the figure, in this example, decoder (15) of decoding terminal (12) has the following parts: an ES decoder that decodes the ES to recover the video data, audio data, etc., and an ES scene description decoder that decodes the scene description, and, at the same time, forms the scene using the video data, audio data, etc., based on the decoded scene description.

[0140]

Here, for the data delivery system in Embodiment 3 having filter (23) in the second example, when the transmittable bandwidth of transmission medium (22) and the traffic jam state change, delay and loss take place in the transmitted data. In order to solve such problem, the following scheme is adopted.

[0141]

Said transmitter (5) of server (10) has a function of attaching a number (coded serial number) to each packet of the data sent out to the transmission line (transmission medium (22)). On the other hand, receiver (13) of decoding terminal (12) has a function of detecting a loss in data (data loss proportion) by monitoring omissions in the serial number (coded serial number) added to each received packet. As another scheme, transmitter (5) of server (10) has a function of adding time information (encoded time information) to the data sent to the transmission line, while receiver (13) of decoding terminal (12) has a function of detecting a delay in transmission from the time information by monitoring the time information (coded time information) attached to the data received from the transmission line. When receiver (13) of decoding terminal (12) detects a data loss proportion of a transmission line or a transmission delay or the like of the

transmission line as mentioned previously, the detection information is sent (reported) to transmitter (5) of server (10).

[0142]

Also, transmitter (5) of server (10) has a function of detecting the transmission state. With said transmission state [detection] function, the transmittable bandwidth of the transmission line and the traffic jam state can be detected based on information of the data loss proportion or transmission delay, etc., of the transmission line sent from receiver (13) of decoding terminal (12). That is, with the transmission state detection function, if a high data loss occurs, it is judged that the transmission line is jammed, or, if the transmission delay increases, it is judged that the transmission line is jammed. Also, when a bandwidth reservation type transmission line is used, the transmission state detection function allows direct detection of the idle bandwidth (transmittable bandwidth) that can be used by server (10). With regard to the transmission bandwidth, when electromagnetic waves or other transmission media depending on the weather conditions, etc., is used, the user may preset the transmission bandwidth in advance corresponding to the weather conditions, etc. The detected information of the transmission state with said transmission state detection function is sent to controller (26) of filter (23).

[0143]

Said controller (26) controls such that switching is selectively executed between different ESs having different bit rates in ES processor (25) based on the detection information of the transmittable bandwidth of the transmission line and the traffic jam state, or, it controls by adjusting the coding bit rate when coding is carried out according to ISO/IEC 13818 (so-called MPEG2) at ES processor (25). That is, when it is detected that the transmission line is jammed, an ES with a lower bit rate is output from ES processor (25). As a result, delay in the data can be avoided.

[0144]

Also, for a system constitution in which an unspecified plurality of decoding terminals (12) are connected to server (10), the specifications of decoding terminals (12) are not pre-unified, and ESs are transmitted from server (10) to said decoding terminals (12) with various processabilities, receiver (13) of decoding terminal (12) has a transmission request processing function. With said transmission request processing function, a transmission request signal for requesting an ES corresponding to the processability of the local decoding terminal (12) is sent to server (10). The transmission request signal also contains a signal that represents the ability of local decoding terminal (12). Examples of signals representing the ability of local

decoding terminal (12) and sent to server (10) with said transmission request processing function include the memory size, the resolution of the display part, the arithmetic and logic operation ability, the buffer size, the coding format of the decodable ES, the number of decodable ESs, the bit rate of the decodable ES, etc. Said transmitter (5) that receives said transmission request signal sends the transmission request signal to controller (26) of filter (23). Said controller (26) controls ES processor (25) so that ESs are transmitted corresponding to the performance of decoding terminal (12). For the video signal transformation processing when an ES is transformed corresponding to the performance of decoding terminal (12) by ES processor (25), for example, a video signal transformation processing method that has been proposed by the present patent applicant can be used.

[0145]

In addition, said controller (26) also controls scene description processor (24) in addition to said ES processor (25), corresponding to the state of the transmission line detected according to the transmission state detection function of transmitter (5). Also, if decoding terminal (12) is a decoding terminal that requests a scene description corresponding to the local decoding and display performance, controller (26) controls ES processor (25) and scene description processor (24) corresponding to a signal representing the local ability of the decoding terminal sent with the transmission request processing function of receiver (13) of decoding terminal (12). Here, said controller (26), scene description processor (24) and ES processor (25) may have an integrated constitution.

[0146]

In the following, the selection method when ES processor (25) selects a prescribed ES to be transmitted from among plural ESs under control of controller (26) will be explained.

[0147]

For each ES of said plural ESs, said controller (26) holds transmission priority information representing the priority for transmission, and it determines the ES that can be transmitted in the descending order of the transmission priority corresponding to the state of the transmission line when ES is transmitted or a request from decoding terminal (12). That is, corresponding to the state of the transmission line or a request from decoding terminal (12) when an ES is transmitted, controller (26) controls ES processor (25) so that the transmittable ES are transmitted in the descending order of the transmission priority. Here, for example, a scheme in which controller (26) holds the transmission priority information has been explained. However, said information may also be stored in storage part (9).

[0148]

Figure 17 shows an example in which various ESs have their transmission priorities when three ESs exist, that is, ESa, ESb, ESc. That is, in the example shown in Figure 17, ESa has a transmission priority of 30, ESb has a transmission priority of 20, and ESc has a transmission priority of 10. For the transmission priority, the smaller the value, the higher the priority for transmission. In Figure 17, Ra represents the transmission bit rate when ESa is transmitted, Rb represents the transmission bit rate when ESb is transmitted, and Rc represents the transmission bit rate when ESc is transmitted.

[0149]

Here, ES processor (24) is controlled by controller (26) so that when the transmittable bit rate R is determined according to the state of the transmission line and a request from decoding terminal (12), ESs are selected and transmitted in the descending order of the transmission priority in the range where said transmittable bit rate R is not exceeded.

[0150]

That is, for example, when the relationship between the transmittable bit rate R and the transmission bit rate of respective ESs is represented by equation (4), controller (26) controls ES processor (25) so that only ESc with the highest transmission priority is selected and transmitted.

[0151]

$$Rc \leqq R < (Rc + Rb) \qquad (4)$$

Also, for example, when the relationship between transmittable bit rate R and the transmission bit rate of respective ESs is represented by equation (5), controller (26) controls ES processor (25) so that ESc with the highest transmission priority and ESb with the next highest transmission priority (the second highest transmission priority) are selected and transmitted.

[0152]

$$(Rc + Rb) \leqq R < (Rc + Rb + Ra) \qquad (5)$$

Also, for example, when the relationship between transmittable bit rate R and the transmission bit rate of respective ESs is represented by equation (6), controller (26) controls ES processor (25) so that all ESs are selected and transmitted.

[0153]

$$(Rc+Rb+Ra) \leqq R \qquad (6)$$

In this way, for the data delivery system in Embodiment 3 having filter (23) of the third example, controller (26) holds the transmission priority information for each ES, and, it determines the transmittable ESs in the descending order of the transmission priority corresponding to the state of the transmission line when ES is transmitted and a request from decoding terminal (12). As a result, the more important ES with higher priority from among plural ESs can be transmitted.

[0154]

In the example explained above, ES selection and scene description transformation are carried out according to a preset priority. However, one may also adopt a scheme in which the priority is changed in company with transformation of said ES. Also, when the priority changes in company with transformation of an ES, the priority may be changed by means of ES processor (25).

[0155]

Figure 18 shows an example of transmission priority transformation by ES processor (25) in company with transformation of the bit rate of ESa to Ra'. In the example shown in Figure 18, the bit rate Ra of ESa is lower than that of Ra in the example shown in Figure 17. When the bit rate decreases, the transmission priority becomes higher with the transformation ("30" in Figure 17 is transformed to "15" in Figure 18).

[0156]

Said transmission priority, in addition to the case in which a preset value is held by controller (26), may also be set corresponding to the ES bit rate, image frame or another coding parameter. For example, as shown in Figure 19, by holding the relationship Ps(R) between transmittable bit rate R of an ES and the transmission priority, the transmission priority also can be set corresponding to the bit rate of the ES. For example, considering that the higher the bit rate, the higher the transmission cost, in the example shown in Figure 19, one may also adopt a scheme in which the transmission priority is lowered for a higher bit rate of an ES, so that an ES with a lower transmission cost (with a lower bit rate) can be transmitted with higher priority.

[0157]

When the ES itself has a clearly defined image frame just like image data, one may also set the transmission priority corresponding to the image frame. For example, as shown in Figure 20, which illustrates an example of relationship Ps(S) between image frame region S of an ES and the transmission priority, by holding said relationship Ps(S) between said image frame region S and the transmission priority, the transmission priority can be set according to the image frame of the ES. That is, considering that a larger image frame usually leads to a higher transmission cost, in the example shown in Figure 20, the larger the image frame, the lower the assigned transmission priority, so an ES believed to have a lower transmission cost is transmitted with a higher priority.

[0158]

As explained above, a method for setting the transmission priority corresponding to the ES bit rate, the image frame, or another coding parameter can also be adopted in the case in which the transmission priority is changed in company with transformation of the ES by ES processor (25). For example, when ES processor (25) transforms the ES with bit rate Ra to bit rate Ra', the transmission priority shown in Figure 19 can be changed to Ps(Ra').

[0159]

Also, a transmission priority may be assigned for each type of ES, such as a moving picture, still picture, text, etc., and for each ES coding format. For example, by always assigning the highest transmission priority to text, text data can always be transmitted with priority even if the transmittable bit rate is restricted by the state of the transmission line or a request from the decoding terminal.

[0160]

Also, the transmission priority may be determined based on the preference of the user. That is, server (10) holds information about the type of ES, such as moving pictures, still pictures, text, or the like favored by the user, the ES coding format, ES coding parameters, and other items that are preferred. As a result, a high transmission priority can be assigned to an ES of an ES type, coding format, coding parameters, etc., preferred by the user. Consequently, even if there is a restriction on the transmittable bit rate corresponding to the state of the transmission line and a requirement from the decoding terminal, an ES can be transmitted with priority and at a high quality corresponding to the preference of the user.

[0161]

As explained above, controller (26) holds transmission priority information for each ES, and it determines the transmittable ESs in descending order of the transmission priorities corresponding to the state of the transmission line for transmission and a request from decoding terminal (12), so the more important ESs can be transmitted with higher priority.

[0162]

As to be explained later, for filter (23) in the third example in Embodiment 3 of the present invention, data meeting the bit rate or another evaluation standard can be delivered even in special playback. That is, scene description processor (24) arranged in filter (23) in said third example can carry out the first through fifth scene description processing schemes to be explained below under control of controller (26).

[0163]

As a first scheme of scene description processing, for example, filter (23) in the third example can output the scene corresponding to the ES output from ES processor (25). That is, under control of controller (26), scene description processor (24) can output a scene description appropriate for the ES output from ES processor (25). In the following, the first scene description processing scheme will be explained in more detail with reference to Figures 21-25.

[0164]

Figure 21 is a diagram illustrating an example of display of a scene composed of a moving picture ES and a still picture ES. In Figure 21, Esi represents a scene display region, Emv represents a moving picture ES display region in scene display region Esi, and Esv represents a display region of a still picture ES in scene display region Esi.

[0165]

As shown in Figure 22, a scene description corresponding to scene display region Esi shown in Figure 21 is represented by the contents and text of description using MPEG4 BIFS.

[0166]

For the scene description shown in Figure 22, two cubes are contained, and both have their surfaces bonded with the textures of a moving picture and still picture by assignment. The objects are assigned for coordinate transformation by means of a Transform node, and, by means of the values of the translation fields indicated by #500 and #502 in the figure (the origin position of the local coordinates), the object is translationally moved and arranged in the scene.

By means of the values indicated as #501 and #503 (scaling of the local coordinates), the object contained in the Transform node is assigned for enlargement/contraction.

[0167]

Here, for example, if it is necessary to decrease the data quantity in transmission in a case in which it is necessary to decrease the bit rate of the delivery data corresponding to the state of the transmission line (transmission medium (22)) or a request from decoding terminal (12), transformation processing of the ESs is carried out so that the bit rate of the moving picture ES that requires a large data quantity in transmission is decreased. For the still picture at this time, for example, a high resolution still picture ES has been transmitted and accumulated at the decoding terminal.

[0168]

In this case, for a data delivery system of the prior art, whether or not the bit rate of the ES is adjusted, the same scene constitution is adopted for decoding and display. Consequently, the image quality degrades significantly for a moving picture with a lower bit rate. Here, a more detailed explanation is possible with reference to the example shown in Figure 21. Here, for the data delivery system in the prior art, even if the bit rate of the moving picture ES is adjusted to be decreased for display in display region Emv of the moving picture ES shown in Figure 21, because decoding and display of the ES (display on display region Emv of a moving picture ES not fitting the actual bit rate) are carried out while the scene constitution is the same as that before adjustment, so the moving picture becomes more imprecise (for example, the spatial resolution becomes lower), and the image quality degrades significantly.

[0169]

On the other hand, when the bit rate of a moving picture ES decreases, for example, as shown in Figure 23, display region Emv of the moving picture ES becomes narrower, but degradation in the image quality of the moving picture displayed on display region Emv of said moving picture ES (degradation in the spatial resolution in this example) may be not significant. Also, in this embodiment, for a still picture, a still picture ES is transmitted and accumulated at the decoding terminal. When the still picture has a high resolution, and display region Esv of still picture ES in Figure 21 is narrow so that it does not match said resolution, as shown in Figure 23, if display region Esv of the still picture ES is widened, it is believed that the resolution can be suitably exploited. In this way, the measures in narrowing the display region Emv of a moving picture ES and extending display region Esv of a still picture ES cannot be taken if the scene description is not changed to a scene description that represents the contents.

[0170]

Here, scene description processor (24) arranged in filter (23) in the third example changes the scene description and outputs it dynamically corresponding to the adjustment of the bit rate of the ES in ES processor (25). In other words, for controller (26) in the third example, when ES processor (25) is controlled to adjust the bit rate of the ES, scene description processor (24) is also controlled so that a scene description matching the ES output from ES processor (25) is output. As a result, the image quality does not degrade significantly when the bit rate of the moving picture is decreased as described in said example. In this example, because the resolution of the transmitted still picture is exploited, as shown in Figure 23, the following measures are taken: display region Emv of the moving picture ES is narrowed, while display region Esv of the still picture ES is extended.

[0171]

In the following, the operation of controller (26) for realizing the aforementioned objective will be explained in more detail with reference to Figure 24.

[0172]

As shown in Figure 24, when the bit rate of the delivery data should be decreased corresponding to the state of the transmission line or a request from decoding terminal (12), controller (26) controls ES processor (25) so that at time T, moving picture ES (293) with a bit rate lower than that of moving picture ES (292) is output.

[0173]

Also, controller (26) controls scene description processor (24) so that at time T, scene description (290) corresponding to scene display region Esi shown in Figure 21 is changed to scene description (291) corresponding to scene display region Esi shown in Figure 23. That is, in this case, under control of controller (26), scene description processor (24) transforms the scene description shown in Figure 22 representing the scene description region Esi shown in Figure 21 to the scene description shown in Figure 25 and representing scene display region Esi shown in Figure 23. Here, just as in the case shown in Figure 22, the scene description shown in Figure 25 displays the text contents of scene description described by means of MPEG4 BIFS.

[0174]

Different from the scene description in Figure 22, in the scene description shown in Figure 25, the values of the translation fields indicated as #600 and #602 (site of origin of the

local coordinates) are changed, so that two cubes are moved, and, by means of the values of the translation fields indicated as #601 and #603 in the figure (scaling of the local coordinates), the cube with a moving picture (Emv in Figure 23) bonded on it is transformed to a smaller size, and, on the other hand, the cube with a still picture (Esv in Figure 23) bonded on its surface is transformed to a larger size.

[0175]

The transformation processing from the scene description shown in Figure 22 to the scene description shown in Figure 25 as in said first scene description processing may be carried out with any of the following processing schemes: in one processing scheme, in scene description processor (24), from plural scene descriptions previously stored in storage part (9), the scene description corresponding to the ES output from ES processor (25) (the scene description shown in Figure 25) is selectively read and delivered; in another processing scheme, the scene description read from storage part (4) (the scene description shown in Figure 22) is transformed to a scene description corresponding to the ES output from ES processor (25) (the scene description shown in Figure 25); in yet another processing scheme, scene description data corresponding to the ES output from ES processor (25) (the scene description shown in Figure 25) are generated or coded and delivered. Also, if a scene description method that allows description of the variation in the scene description is adopted, only the variation is transmitted. In said example, a case in which display region Emv of a moving picture ES is narrowed when the bit rate of the moving picture ES is decreased was explained. Conversely, when display region Emv of a moving picture ES is extended as the bit rate is increased, the scene description transformation of the present invention can obviously be adopted. In addition, in said example, a case in which a high resolution still picture ES is transmitted and stored in advance was explained. For example, when the still picture transmitted and stored in advance has a low resolution, a still picture ES with a high resolution is newly transmitted, and the corresponding scene description is transmitted. In addition, in this embodiment, a moving picture and a still picture are taken as examples, but the present invention may also be adopted in a case in which the scene description is changed corresponding to adjustment of the bit rate of other multimedia data.

[0176]

According to the first scene description processing scheme explained with reference to Figures 21-25, by means of processing for transformation of the scene description that represents the scene constitution information, the scene description can be transmitted corresponding to the state of the transmission line or a request from decoding terminal (12). For example, when

transformation of an ES is carried out by means of ES processor (25), an optimum scene description can be transmitted to the ES after transformation.

[0177]

In the following, a second scene description processing scheme will be explained.

[0178]

For example, when the bit rate of an ES from ES processor (25) is transformed and the information for decoding the ES is changed as needed corresponding to the state of the transmission line and decoding terminal (12), as scene description processing scheme, filter (23) carries out scene description processing by transforming and transmitting the scene description itself containing the information needed for decoding said ES, so the information needed for decoding does not have to be extracted from the data of the ES itself at the signal terminal. That is, under control of controller (26), said scene description processor (24) carries out ES transformation processing by means of ES processor (25), and the information needed for decoding said ES is changed. In this case, a scene description containing the information needed for decoding of ES can be output. Here, the information needed for decoding ES includes the coding format of ES, the buffer size needed for decoding, the bit rate, etc. In the following, the second scene description processing scheme will be explained in more detail with reference to said figures and Figures 26 and 27.

[0179]

As shown in Figure 26, an example of information needed for decoding of the ES used in the scene for explanation with reference to said Figures 21 and 22 is described by means of descriptor Object Descriptor defined in MPEG4. In the scene description shown in Figure 22, the moving picture mapped as texture on the surface of the object is assigned by value 3 (= url3). This corresponds to 0Did = 3 as the identifier of the Object Descriptor shown in Figure 26. ES_Descriptor contained in the Object Descriptor of identifier 0Did = 3 describes information pertaining to the ES. Also, in the figure, ES_ID is the identifier that uniquely specifies the ES. This identifier ES_ID is further correlated with the identifier, port number, etc., of the header in the transmission protocol for transmission of the ES, so it actually corresponds to the actual ES.

[0180]

Also, the description of ES_Descriptor contains the information needed for decoding the ES known as Decoder Config Descriptor. The information of said descriptor Decoder Config

Descriptor includes the buffer size, the maximum bit rate, the average bit rate, etc., needed for decoding of the ES.

[0181]

On the other hand, as shown in Figure 27, an example of the information needed for decoding of an ES in company with the scene description after transformation processing in scene description processor (24) corresponding to the scene shown in Figure 23 as mentioned previously is described by means of descriptor Object Descriptor defined by MPEG4. The decoding buffer size (buffer Size DB), maximum bit rate (max Bit Max) and average bit rate (avg Bit Rate) of a moving picture (with 0Did of 3, referenced from the scene description) that changes due to transformation of an ES are transformed as shown in Figure 27 from the description in Object Descriptor shown in Figure 26 before said transformation. That is, in the example shown in Figure 26, one has bufferSizeDB = 4000, maxBitRate = 1000000, avgBitRate-1000000. In Figure 27, they are transformed to bufferSizeDB = 2000, maxBitRate = 5000000, avgBitRate-5000000.

[0182]

The transformation processing of the information needed for decoding of an ES in company with the scene description as in said second scene description processing may be carried out with any of the following processing schemes: in one processing scheme, in scene description processor (24), from the plural information items needed for decoding of an ES and previously stored in storage part (9), the information corresponding to the ES output from ES processor (25) (the information shown in Figure 27) is selectively read and delivered; in another processing scheme, the information needed for decoding of an ES (the information shown in Figure 26) and read from storage part (9) is transformed into information needed for decoding of the ES output from ES processor (25) (the information shown in Figure 27) and is output; in yet another processing scheme, the information needed for decoding an ES (the information shown in Figure 27) output from ES processor (25) is encoded and delivered.

[0183]

According to the second scene description processing scheme explained above, if the information needed for decoding of an ES is changed by transforming the bit rate or the like of the ES corresponding to the state of the transmission line or a request from decoding terminal (12), as shown in Figure 27, by changing the information needed for decoding of the ES contained in the scene description and sending it to decoding terminal (12), the necessity of

extraction of information needed for decoding of the ES from the data of the ES itself at decoding terminal (12) can be avoided.

[0184]

In the following, a third scene description processing scheme will be explained.

[0185]

According to the third scene description processing scheme, for filter (23), since the scene description is transformed and output explicitly so that the number of ESs that form the scene is adjusted, only an ES matching the transmission bandwidth can be transmitted. On the other hand, at decoding terminal (12), the ES needed for display, etc., can be judged independently of a delay in the arrival of the ES data or loss of the data. That is, in this example, under control of controller (26), scene description processor (24) transforms and outputs the scene description explicitly so that the number of ESs is adjusted, and the scene description decoding function arranged at decoder (15) of decoding terminal (12) is judged independently of a delay in the arrival of ES data and a loss in data of the ES needed for display, etc. In the following, the third scene description processing scheme will be explained more specifically with reference to aforementioned figures and Figures 28 and 29.

[0186]

In Figure 28, a scene description in the case of deletion of the moving picture ES from the scene explained with reference to Figures 21 and 22 is described by means of MPEG4 BIFS (described as text that can be easily understood). Figure 29 shows an example of a scene displayed based on the scene description shown in Figure 28. In this case, only image ES display region Eim (such as a still picture ES display region) is arranged in scene display region Esi. For the ES used in the scene description in Figure 28, because it is possible to judge the fact that 0Did is only the ES of 4 for use in the scene description in Figure 28 from the scene description, at decoding terminal (12), it is possible to judge that it does not depend on a delay in the arrival of the ES data and loss in the data even if a moving picture ES with an 0Did of 3 did not arrive. In addition, as in the example shown in Figures 26 and 27, by deleting the description of Object Descriptor with an 0Did of 3, it is possible to judge that the moving picture ES with an 0Did of 3 is not needed.

[0187]

In the example of the third scene description processing scheme, when a request for transmission, in which the processing load for decoding and forming the scene is temporarily

decreased, is transmitted from decoding terminal (12), in filter (23), for example, the scene description shown in Figure 22 is changed to the scene description shown in Figure 28, so processing for mapping the moving picture as texture to the scene is explicitly not needed, and this is notified to decoding terminal (12). As a result, at decoding terminal (12), it is possible to decrease the processing load for decoding the scene.

[0188]

The transformation processing from the scene description shown in Figure 22 to the scene description shown in Figure 28 as in said third scene description processing scheme may be carried out with any of the following processing schemes: in one processing scheme, in scene description processor (24), the scene description (the scene description shown in Figure 28) corresponding to the ES number output from ES processor (25) is selectively read and output from the plural scene descriptions previously prepared in storage part (9); in another processing scheme, the scene read from storage part (9) is taken as input, and transformation and output processing is carried out to the scene description (scene description in Figure 28) with the portion of data corresponding to ESs not output (the data in the scene description) deleted; in yet another processing scheme, when the scene description is coded and output, processing is carried out without coding for the portion corresponding to ESs not output.

[0189]

As explained above, according to the third scene description processing scheme, with the aforementioned transformation of the scene description, at server (10), the demanded scene can be recovered at decoding terminal (12) at a desired timing. Also, according to the third scene description processing scheme, in scene description processor (24), it is possible to sequentially delete portions of data with lower importance in a scene description until the transmission bandwidth and the processing performance of decoding terminal (12) are matched. Also, by means of the third scene description processing scheme, when a margin is generated in the processing performance of decoding terminal (12), a more detailed scene description can be sent, and, as a result, an optimum scene with respect to the processing performance can be decoded and displayed at decoding terminal (12).

[0190]

In the following, a fourth scene description processing will be explained.

[0191]

As the fourth scene description processing scheme, at server (10) of this embodiment, the complexity of the scene description is transformed corresponding to the state of the transmission line or a request from decoding terminal (12), so the data quantity of the scene description can be adjusted, and the processing load on decoding terminal (12) can be adjusted. That is, in this example, under control of controller (26), scene description processor (24) adjusts and outputs the data quantity of the scene description corresponding to the state of the transmission line or a request from decoding terminal (12). In the following, the fourth scene description processing scheme will be explained more specifically with reference to Figures 30-33.

[0192]

Figure 30 shows a case in which the scene description for displaying an object described by polygons is described according to MPEG4 BIFS (described as text to facilitate understanding). In the example shown in Figure 30, in order to simplify the explanation, the coordinates of the polygons are not shown. Also, in the scene description shown in Figure 30, Indexed Face Set represents the geometric object for which the apex coordinates assigned with points in Coordinates can be connected in the sequence assigned with CoordIndex. In addition, Figure 31 shows an example of display of the scene represented by decoding the scene description in Figure 30 (example of display of polygonal object).

[0193]

In the example of the fourth scene description processing, depending on the state of the transmission line, for example, when the data quantity of transmission by server (10) is to be decreased, or when a transmission request for a lower processing load is transmitted from decoding terminal (12), scene description processor (24) of filter (23) transforms the scene description to a simpler scene description. For example, in the example of scene description shown in Figure 32, Indexed Face Set representing the polygons shown in Figure 31 is replaced by a Sphere representing a ball as shown in Figure 33, so the data quantity of the scene description itself can be decreased, and the processing load for forming the scene can be decreased at decoding terminal (12). That is, for the polygons shown in Figure 31, the various values for representing the polyhedron are needed. On the other hand, for the ball shown in Figure 33, they are not needed. Consequently, the data quantity for scene description can be decreased. Also, at decoding terminal (12), the complicated processing for displaying the polyhedron becomes the simple processing for representing the ball, so that the load on processing becomes lower.

[0194]

Transformation processing from the scene description shown in Figure 30 to the scene description shown in Figure 32, such as said fourth scene description processing, may be carried out with the following schemes: according to one scheme, in scene description processor (24), from the plural scene descriptions previously prepared in storage part (9), the scene description that meets the evaluation standard matching the state of the transmission line or a request from decoding terminal (12) is selected and output; according to another scheme, a scene description read from storage part (9) is input, and it is transformed to a scene description meeting said evaluation standard; and, according to yet another scheme, a scene description meeting said evaluation standard is coded and output. Here, as said evaluation standard, the standard may represent the complexity of the scene description, such as the data quantity of the scene description, the number of nodes or polygons, etc.

[0195]

Other processing schemes for transforming the complexity of the scene description in scene description processor (24) include processing in which a complex portion of data is replaced by a simpler portion of data as shown in Figure 32 or the reverse processing, processing in which a portion of data are removed or the reverse processing, processing in which the quantization step is changed to adjust the data quantity of the scene description data when the scene description is coded; etc. Also, control of the data quantity of scene description by adjusting the quantization step in coding can be realized as follows. For example, by means of MPEG4 BIFS, for each of the coordinates, rotating axis and angle, size, or other quantization category, quantization parameters representing yes/no of the use of quantization and application bit number can be set, and the quantization parameters can be changed even during one substrate description, so the bit number for use in quantization can be decreased. Consequently, the data quantity of the scene description can be decreased.

[0196]

As explained above, according to the fourth scene description processing scheme, by transforming the scene description, a simplified scene demanded by server (10) can be recovered at decoding terminal (12). By means of the fourth scene description processing, in scene description processor (24), the data portions with lower importance in the scene description can be sequentially cut off until the transmission bandwidth or the processability of decoding terminal (12) can be matched.

[0197]

In the following, a fifth scene description processing will be explained.

[0198]

According to the fifth scene description processing scheme, at server (10), by dividing the scene description into plural decoding units corresponding to the state of the transmission line or a request from decoding terminal (12), the bit rate of the scene description data is adjusted, and concentration of a localized processing load at decoding terminal (12) can be avoided. That is, under control of controller (26) and corresponding to the state of the transmission line or a request from decoding terminal (12), scene description processor (24) divides the scene description into plural decoding units, and the scene description timing of the divided decoding units is adjusted and the units are then output. Here, the decoding unit of scene description for decoding at a certain time is the same as the AU of said coding unit. In the following, the fifth scene description processing will be explained more specifically with reference to Figures 34-38.

[0199]

In the example shown in Figure 34, a scene description representing four objects, such as a ball, cube, cone and cylinder, is described by means of one AU of MPEG4 BIFS. Figure 35 shows an example of display of the scene obtained by decoding and displaying the scene description in Figure 34. It shows four objects, that is, ball (41), cube (42), cone (44) and cylinder (43). The scene described in one AU shown in Figure 34 is entirely decoded at an assigned decoding time, and it has to be reflected in the display at an assigned display time. Here, the decoding time (the time required to decode the AU) is called a DTS (Decoding Time Stamp) in MPEG4.

[0200]

In the example of the fifth scene description processing scheme, when the bit rate of the transmitted data is to be decreased corresponding to the state of the transmission line or a request from decoding terminal (12), or when the localized processing load at decoding terminal (12) should be decreased, scene description processor (24) of filter (23) divides the scene description into plural AUs, and the DTS of each AU is offset, so that the localized bit rate for scene description is adjusted to a bit rate corresponding to the state of the transmission line or the request from decoding terminal (12), and the processing quantity needed for decoding processing for each DTS is adjusted to a processing quantity meeting the demand from decoding terminal (12).

[0201]

That is, first, scene description processor (24) divides the scene description shown in Figure 34 into the four of AU1-AU4 shown in Figure 36. Here, first AU1 has an ID of 1 allotted to the Group node that carries out grouping, and the fact that referencing is possible from a succeeding AU is described. By means of MPEG4 BIFS, a portion of a scene can be sequentially added to the grouping node that can be used as a reference. Second AU2 through fourth AU4 describe a command added to the children field of the Group node with an ID of 1 defining that the portion of a scene is the first AU1.

[0202]

Then, for said first AU1 to fourth AU4, scene description processor (24) offsets and assigns DTSs as shown in Figure 37. That is, first DTS1 is assigned for first AU1; second DTS2 is assigned for second AU2; third DTS3 is assigned for third AU3; and fourth DTS4 is assigned for fourth AU4. As a result, the bit rate of the localized scene description data from server (10) to decoding terminal (12) is decreased, and the load of the localized decoding processing generated for each DTS at decoding terminal (12) is decreased.

[0203]

As shown in Figure 38, an object is added to each DTS for the scene obtained by decoding and displaying the scene description divided into four parts as shown in Figure 36 by means of DTS1 to DTS4, and, in the last DTS4, the same scene as that in Figure 35 is obtained. That is, in DTS1, ball (41) is displayed; in DTS2, cube (42) is added; in DTS3, cone (44) is added; and, in DTS4, cylinder (43) is added. As a result, four objects are ultimately displayed.

[0204]

For transformation processing from the scene description shown in Figure 34 to the scene description shown in Figure 36 as the fifth scene description processing, the following schemes may be adopted: as one scheme, at scene description processor (24), from the plural scene descriptions previously prepared in storage part (9), a scene description that meets the evaluation standard corresponding to the state of the transmission line or a request from decoding terminal (12) is selected and output; as another scheme, a scene description read from storage part (9) is taken as input, and transformation is made to the divided scene descriptions (AU1-AU4) until said evaluation standard is met; and, as yet another scheme, the divided scene descriptions (AU1-AU4) are coded and output until said evaluation standard is met. Here, as the aforementioned evaluation standard in the fifth scene description processing, the following standards

representing the limit of scenes contained in each AU, such as the data quantity of one AU, the number of nodes, the number of objects, the number of polygons, etc., contained in one AU, may be adopted.

[0205]

As explained above, according to the fifth scene description processing scheme, the scene description is divided into plural AUs, and the DTS distance is adjusted for each AU, so the average bit rate of the scene description can be controlled, and the load of the localized decoding processing of decoding terminal (12) can be lightened. Here, the average bit rate can be computed by dividing the sum of the data quantity of an AU having a DTS contained in a certain time interval by said time interval. Consequently, scene description processor (24) can adjust the distance of the DTS so that the average bit rate corresponding to the state of the transmission line or a request from decoding terminal (12) can be realized. Also, in said example, a case of dividing an AU was explained. However, one may also adopt a reverse scheme in which plural AUs are combined.

[0206]

In said explanation, the first through fifth scene description processing schemes are adopted individually. However, the various scene description processing schemes may be combined in any combination, so that plural scene description processing schemes are carried out simultaneously. In this case, the combined scene description processing schemes can be realized such that the aforementioned operations and effects can be simultaneously realized.

[0207]

In this embodiment, as an example of scene description, MPEG4 BIFS has been described. However, the present invention is not limited to the aforementioned scheme. For example, when a scene description method that allows description of variation in the scene description is used, the present invention may also be adopted in a case in which only the variation portion is transmitted.

[0208]

In addition, embodiment of the present invention may be carried out by means of either hardware or software.

[0209]

Also, in said explanation, as examples of the scene description, HTML and MPEG4 BIFS have been presented. In addition, VRML, Java (trademark), etc., can be adopted in all of the scene description methods.

[0210]

The present invention is effective for coding of all types of data, such as video data, audio data, still picture data, text data, graphics data, scene description data, etc. In addition, the present invention can be realized by means of either hardware or software.

[0211]

Effect of the invention

According to the present invention, when normal playback is carried out on the receiver side, the data for said normal playback are output, and, when special playback is carried out on the receiver side, time information related to playback of the coding units of the data used in normal playback is transformed and output corresponding to the special playback. As a result, when special playback is carried out on the receiver side, decoding and display for data other than video data are possible. Also, scene description data can be delivered and decoded, and, in addition, a synchronization relationship can be held between data, and data that meet the transmission bit rate or another evaluation standard can be delivered.

Brief description of the figures

Figure 1 is a block diagram illustrating an example of constitution of the data delivery system in an embodiment of the present invention.

Figure 2 is a block diagram illustrating in detail the server of the data delivery system in Embodiment 1.

Figure 3 is a diagram illustrating transformation processing of time information when FFWD play is carried out in Embodiment 1.

Figure 4 is a diagram illustrating transformation processing of time information when slow play is carried out in Embodiment 1.

Figure 5 is a diagram illustrating transformation processing of time information when jump is carried out in Embodiment 1.

Figure 6 is a block diagram illustrating in detail the constitution of a scene of the data delivery system in Embodiment 2.

Figure 7 is a diagram illustrating transformation processing of time information when FFWD play is carried out in Embodiment 2.

Figure 8 is a diagram illustrating variation in the bit rate when FFWD play is carried out in Embodiment 2.

Figure 9 is a block diagram illustrating in detail the constitution of the server of the data delivery system in Embodiment 3.

Figure 10 is a flow chart illustrating the dividing processing in the filter of a first example in Embodiment 3.

Figure 11 is a diagram illustrating a dividing candidate of scene description by means of MPEG4 BIFS in the filter in the first example.

Figure 12 is a diagram illustrating the structure of the scene description in Figure 11.

Figure 13 is a diagram illustrating the result of decoding and display of the scene description shown in Figure 11.

Figure 14 is a diagram illustrating the result of transformation of the scene description shown in Figure 11.

Figure 15 is a diagram illustrating a different transformation candidate of scene description shown in Figure 11.

Figure 16 is a block diagram illustrating the detailed constitution of the filter in the second example of Embodiment 3.

Figure 17 is a diagram illustrating the relationship between the transmission priority, the bit rate and the three ESs in the filter in the second example.

Figure 18 is a diagram illustrating change in the bit rate and change in the transmission priority.

Figure 19 is a diagram illustrating relationship Ps(R) between bit rate R of an ES and the transmission priority.

Figure 20 is a diagram illustrating relationship Ps(S) between image frame region S of an ES and the transmission priority.

Figure 21 is a diagram illustrating the scene display result by means of scene description before transformation in a first scene description processing.

Figure 22 is a diagram illustrating an example of scene description (MPEG4 BIFS) corresponding to the scene shown in Figure 21.

Figure 23 is a diagram illustrating the results of scene display by means of scene description after transformation in the first scene description processing.

Figure 24 is a diagram illustrating the timing of ES transformation and scene transformation in the first scene description processing.

Figure 25 is a diagram illustrating an example of scene description (MPEG4 BIFS) corresponding to the scene shown in Figure 23.

Figure 26 is a diagram illustrating an example of information (MPEG4 Object Descriptor) in company with the scene description shown in Figure 22 needed for decoding of an ES corresponding to the scene shown in Figure 21.

Figure 27 is a diagram illustrating an example of information (MPEG4 Object Descriptor) in company with the scene description shown in Figure 25 needed for decoding of an ES corresponding to the scene shown in Figure 23.

Figure 28 is a diagram illustrating an example of scene description (MPEG4 BIFS) when the Es of a moving picture is deleted from the scene explained with reference to Figures 21 and 22.

Figure 29 is a diagram illustrating the results of display by means of the scene description shown in Figure 28.

Figure 30 is a diagram illustrating an example of scene description (MPEG4 BIFS) for displaying an object described with polygons.

Figure 31 is a diagram illustrating the results of display by means of the scene description shown in Figure 30.

Figure 32 is a diagram illustrating an example of scene description (MPEG4 BIFS) for an object represented by a ball instead of polygons.

Figure 33 is a diagram illustrating the results of display by means of the scene description shown in Figure 32.

Figure 34 is a diagram illustrating an example of scene description (MPEG4 BIFS) composed of four objects.

Figure 35 is a diagram illustrating the results of display by means of the scene description shown in Figure 34.

Figure 36 is a diagram illustrating an example of each scene description (MPEG4 BIFS) for the four AUs obtained by dividing the scene description shown in Figure 34.

Figure 37 is a diagram illustrating the timing of decoding of each AU shown in Figure 36.

Figure 38 is a diagram illustrating the results of display by means of the scene description of each AU shown in Figure 36.

Figure 39 is a block diagram illustrating schematically the constitution of a data delivery system of the prior art.

Figure 40 is a block diagram illustrating schematically the constitution of a data delivery system that eliminates the disadvantage of the data delivery system shown in Figure 39.

Figure 41 is a diagram that gives a brief account of an example of operation (FFWD play) of the data transformation part for video data in the data delivery system shown in Figure 40.

Figure 42 is a diagram that gives a brief account of an example (REWIND play) of the data transformation part for video data in the data delivery system shown in Figure 40.

Figure 43 is a diagram illustrating scene description using VRML and MPEG4 BIFS.

Explanation of symbols

1	Special playback controller
7	Data transformation part
4	Multiplexer
5	Transmitter
9	Storage part
10	Server
12	Decoding terminal
13	Receiver
14	Separator
15	Decoder
16	Scene synthesizer
17	Read part
18	Scheduler
19	Time information rewrite part
23	Filter
24	Scene description processor
25	ES processor
26	Controller

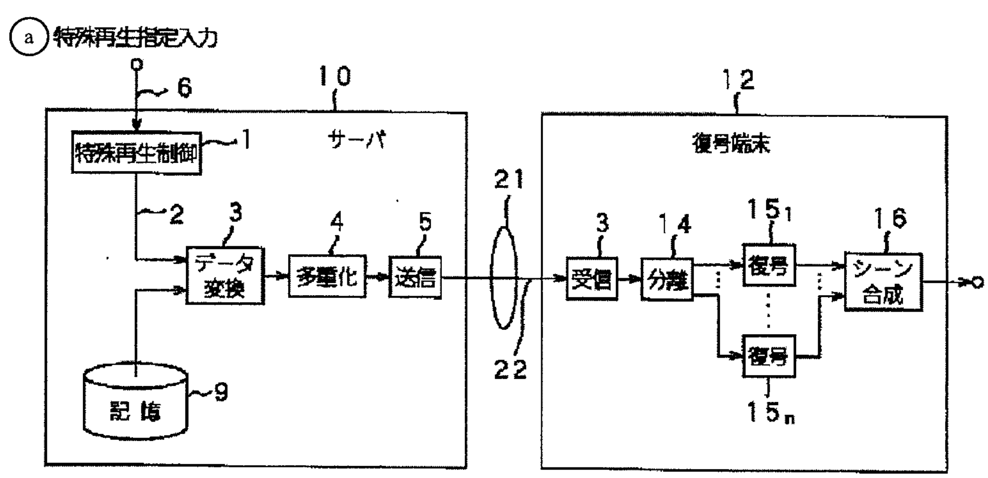


Figure 1

Key:
- a — Special playback assignment input
- 1 — Special playback control
- 3 — Data transformation
- 4 — Multiplexing
- 5 — Sending
- 9 — Storage
- 10 — Server
- 12 — Decoding terminal
- 3 [sic; 13] — Receiving
- 14 — Separation
- $15_1 \ldots 15_n$ — Decoding
- 16 — Scene synthesis

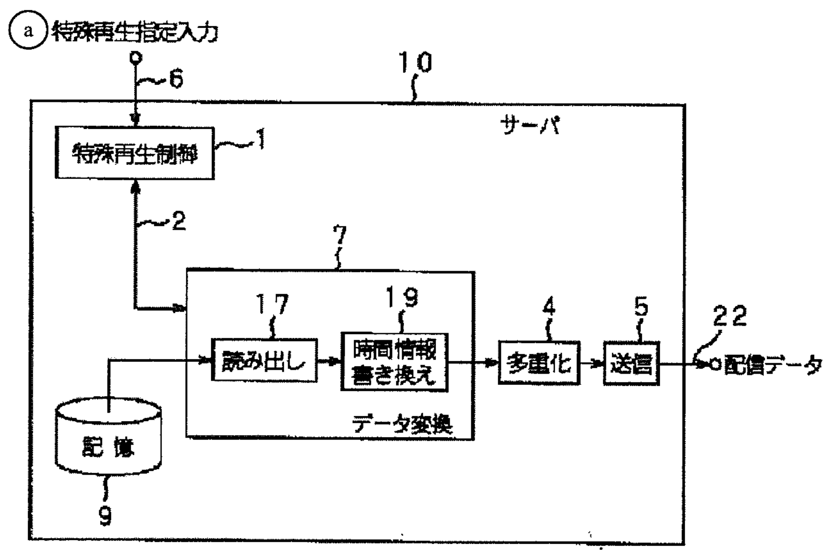


Figure 2

Key:
- a — Special playback assignment input
- 1 — Special playback control
- 4 — Multiplexing
- 5 — Sending
- 7 — Data transformation
- 9 — Storage
- 10 — Server
- 17 — Read
- 19 — Time information rewrite
- 22 — Delivery data

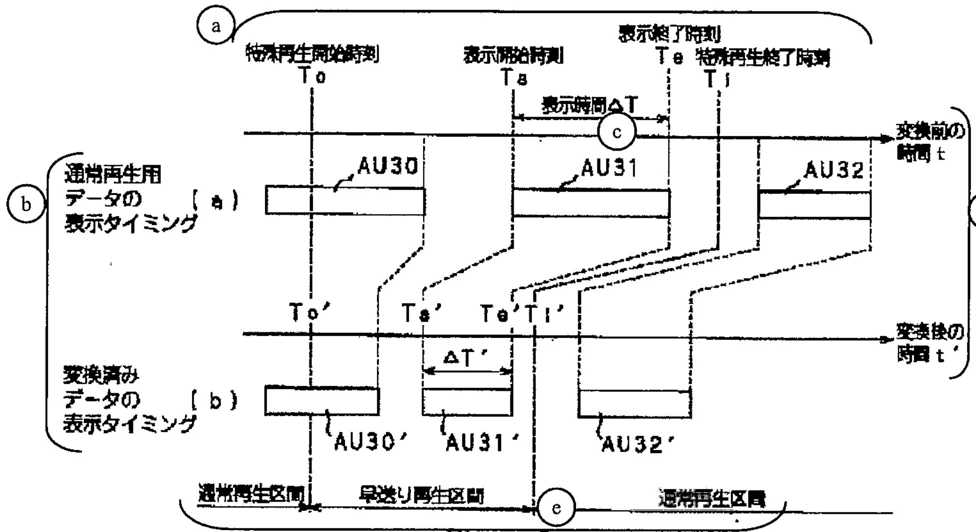


Figure 3

Key:
- a — Special playback start time To
  - Display start time Ts
  - Display end time Te
  - Special playback end time Ti
- b — Display timing of data for normal playback
  - Display timing of transformed data
- c — Display time ΔT
- d — Before-transformation time t
  - After-transformation time t'
- e — Normal playback interval
  - FFWD play interval
  - Normal playback interval

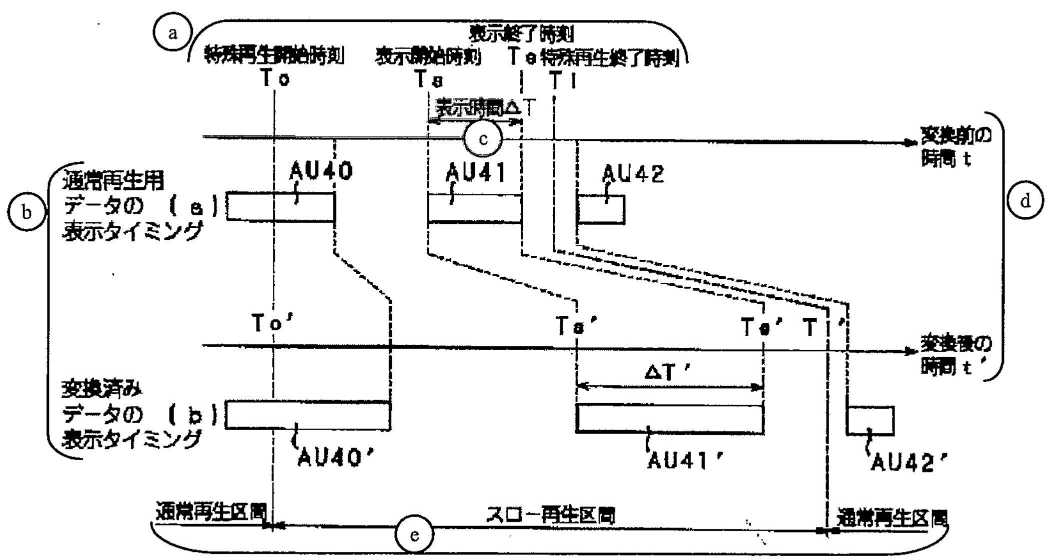


Figure 4

Key:	a	Special playback start time To
		Display start time Ts
		Display end time Te
		Special playback end time Ti
	b	Display timing of data for normal playback
		Display timing of transformed data
	c	Display time ΔT
	d	Before-transformation time t
		After-transformation time t'
	e	Normal playback interval
		Slow play interval
		Normal playback interval

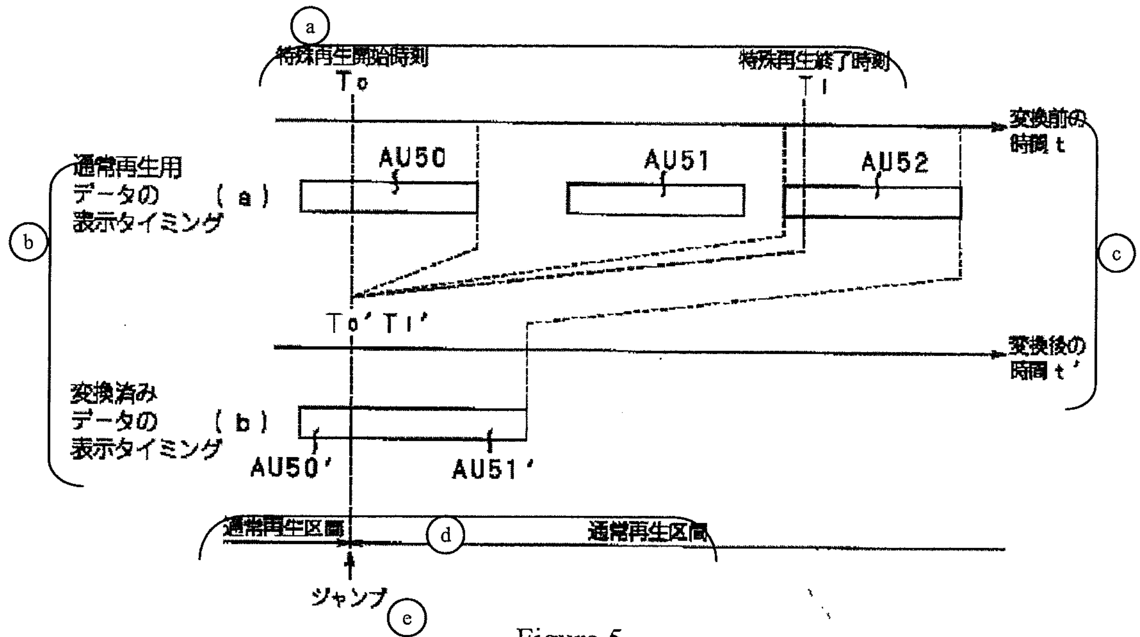


Figure 5

Key: a Special playback start time To
Special playback end time Ti
b Display timing of data for normal playback
Display timing of transformed data
c Before-transformation time t
After-transformation time t'
d Normal playback interval
Normal playback interval
e Jump

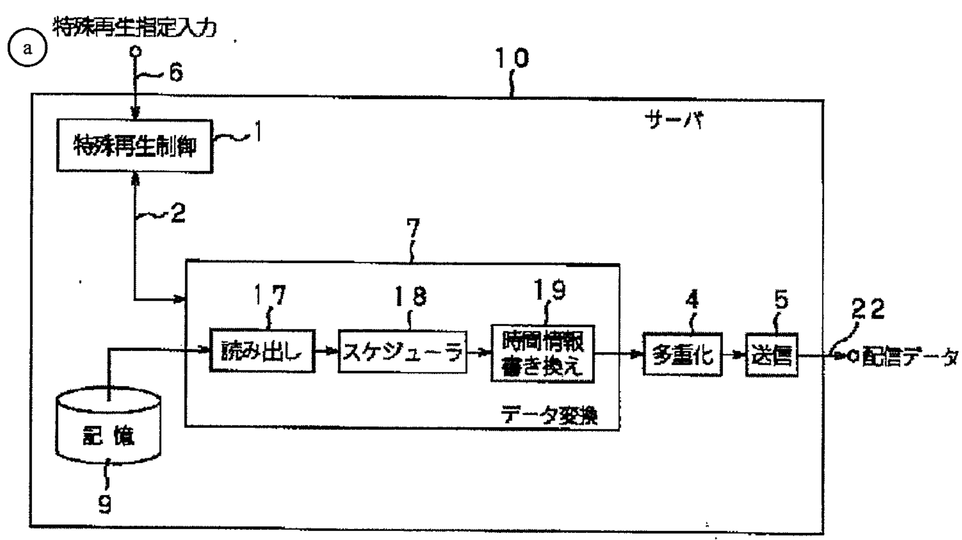


Figure 6

Key: a Special playback assignment input
1 Special playback control
4 Multiplexing
5 Transmission
7 Data transformation
9 Storage
10 Server
17 Read
18 Scheduler
19 Time information rewrite part
22 Delivery data

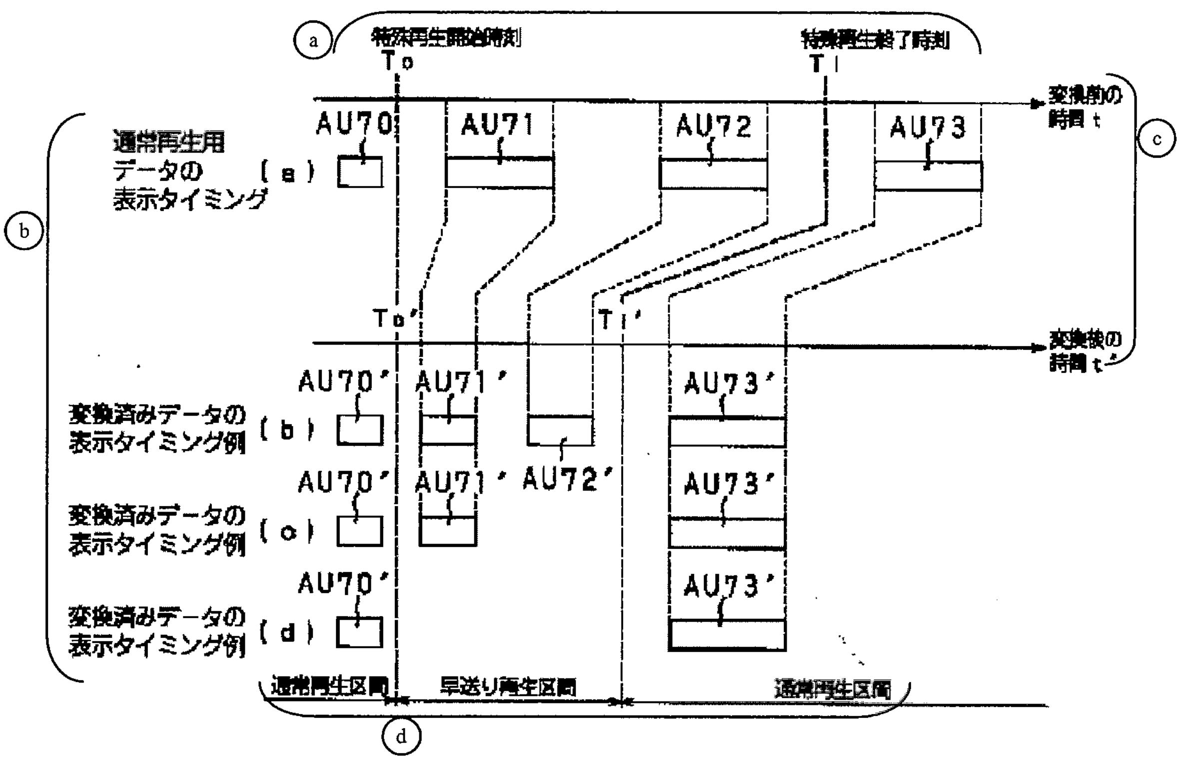


Figure 7

Key: a Special playback start time To
Special playback end time Ti
b Display timing of data for normal playback
Example of display timing of transformed data
Example of display timing of transformed data
Example of display timing of transformed data
c Before-transformation time t
After-transformation time t'
d Normal playback interval
FFWD play interval
Normal playback interval

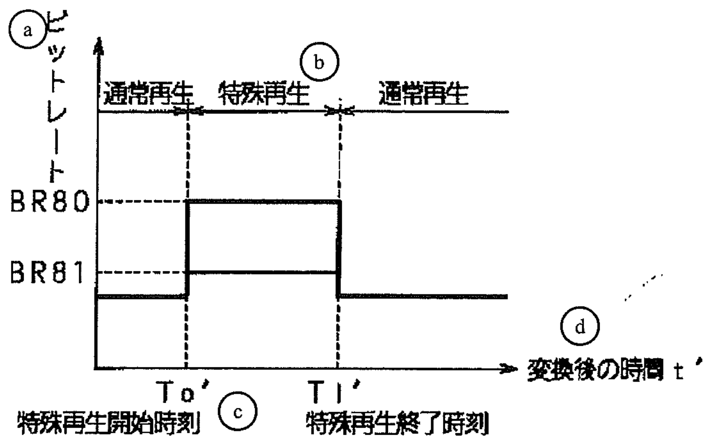


Figure 8

Key:
- a Bit rate
- b Normal playback
  Special playback
  Normal playback
- c Special playback start time
  Special playback end time
- d After-transformation time t'

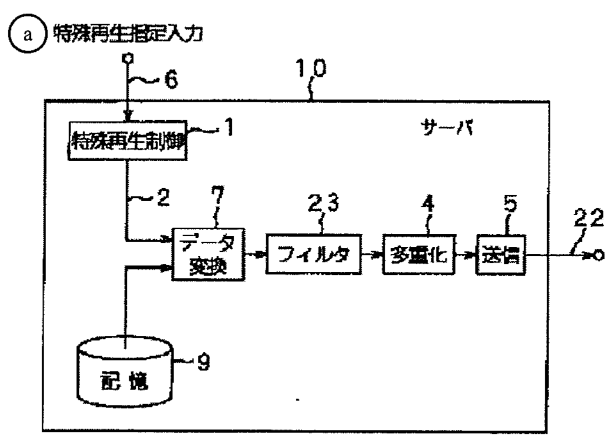


Figure 9

Key:
- a Special playback assignment input
- 1 Special playback control
- 4 Multiplexing
- 5 Transmission
- 7 Data transformation
- 9 Storage
- 10 Server
- 23 Filter

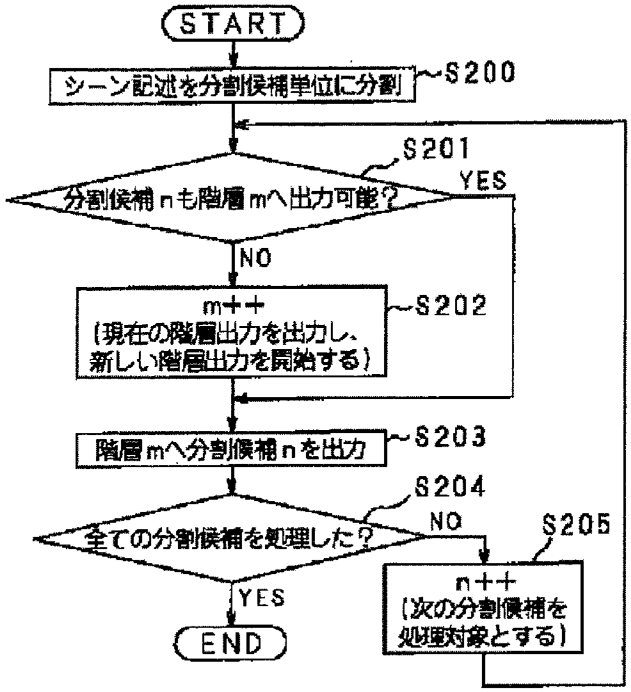


Figure 10

Key:

Key:	S200	Dividing of scene description into division candidate units
	S201	Is it possible to output division candidate n to layer m?
	S202	m++ (output of current layer output, start of new layer output)
	S203	Output of division candidate n to layer m
	S204	Have all of the division candidates been processed?
	S205	n++ (next division candidate is taken as processing object)

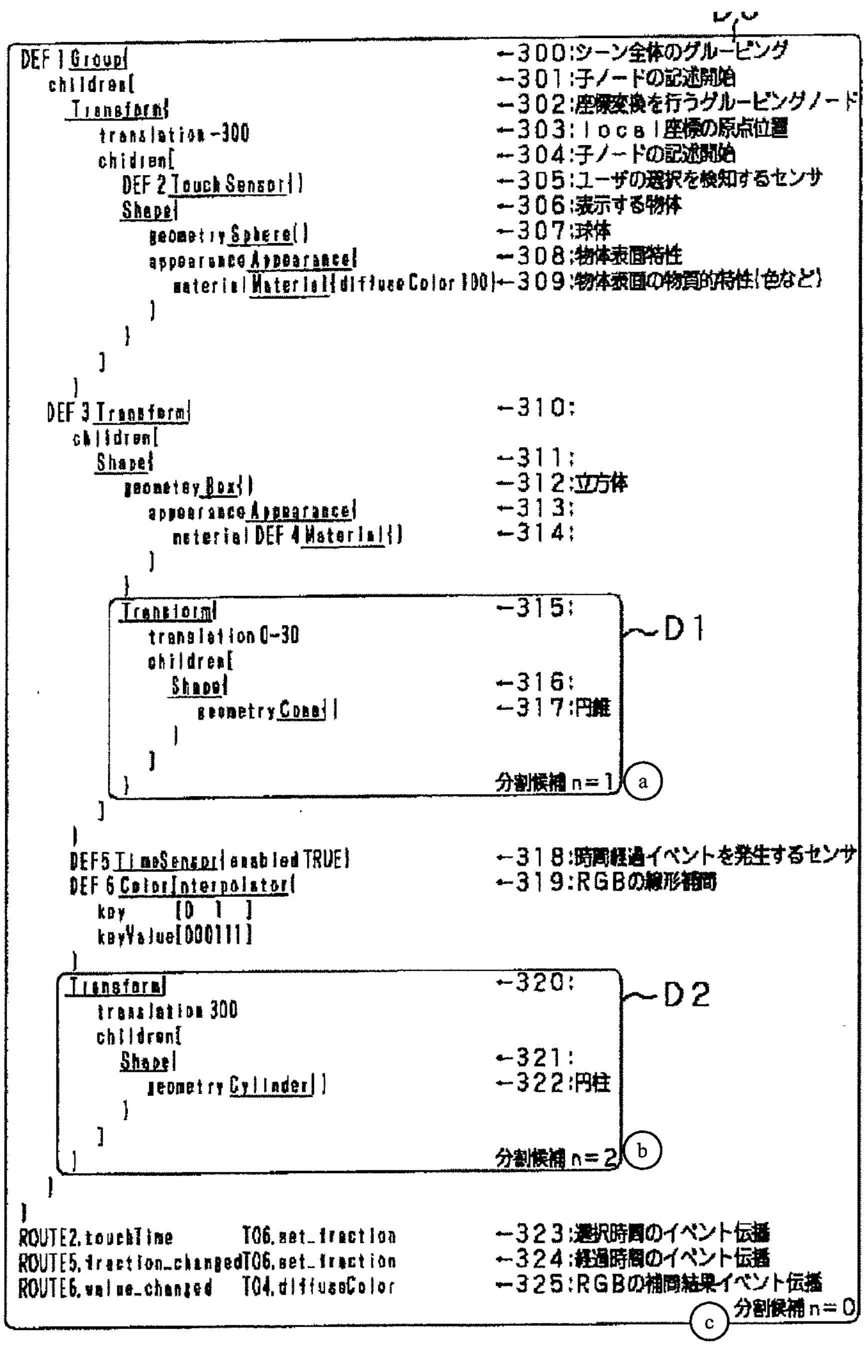


Figure 11

Key:

Key:	a	Division candidate n = 1
	b	Division candidate n = 2
	c	Division candidate n = 0
	300	Grouping of overall scene
	301	Start of description of children node
	302	Grouping node carrying out coordinate transformation
	303	Site of origin of local coordinates
	304	Start of description of children node

305 Sensor detecting selection by the user
306 Displayed object
307 Ball
308 Object surface characteristics
309 Characteristics of substance of object surface (color, etc.)
312 Cube
317 Cone
318 Sensor that generates time course event
319 Linear interpolation of RGB
322 Cylinder
323 Event propagation of selection time
324 Event propagation of lapsed time
325 Event propagation of interpolation result of RGB

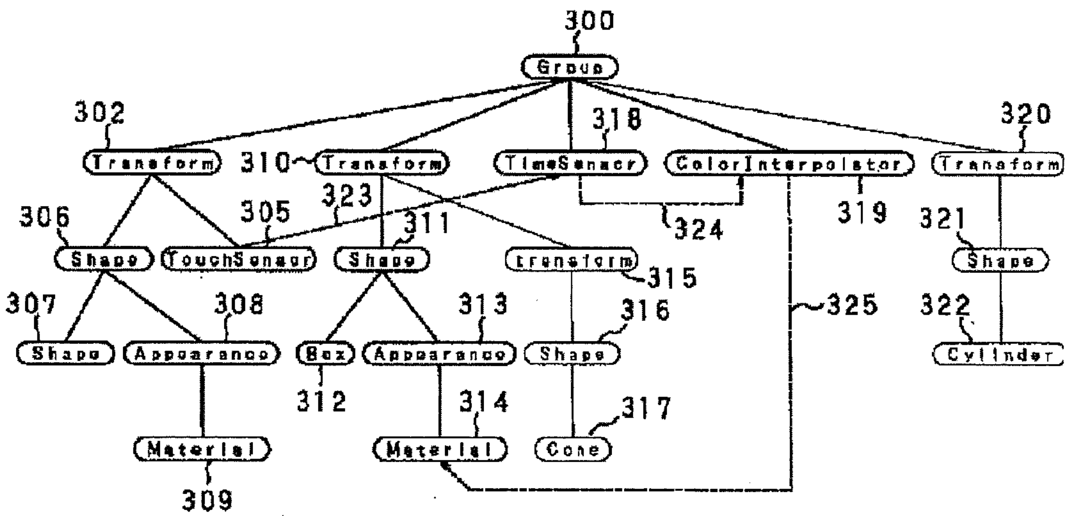


Figure 12

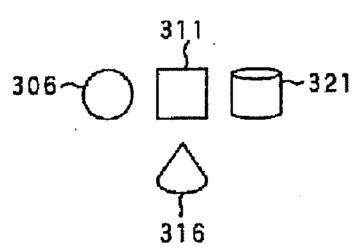


Figure 13

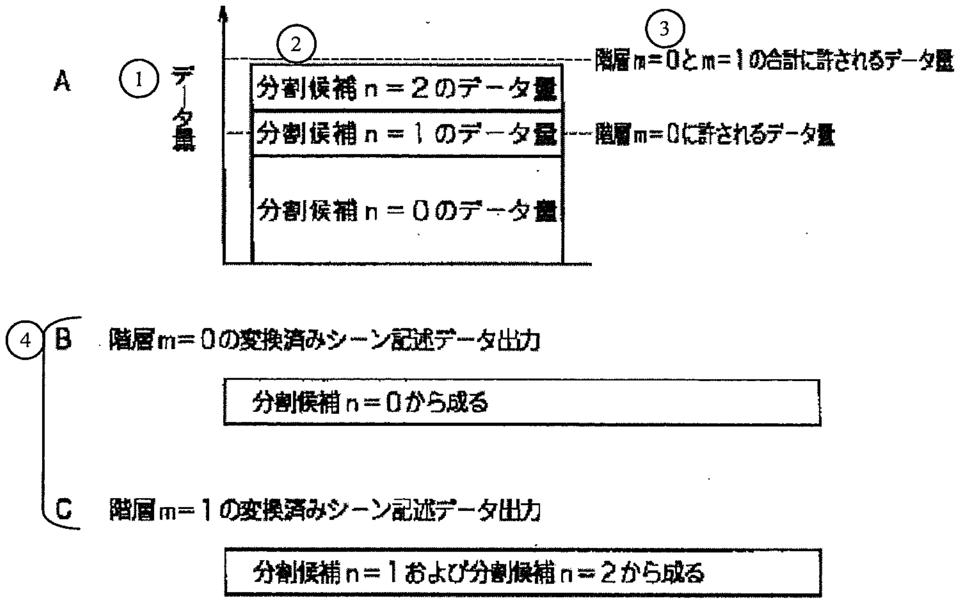


Figure 14

Key: 1 Data quantity

2 Data quantity of division candidate n = 2

Data quantity of division candidate n = 1

Data quantity of division candidate n = 0

3 Data quantity allowed for the sum of layers m = 0, 1

Data quantity allowed for layer m = 0

4 B Transformed scene description data output of layer m = 0

Comprised of division candidate n = 0

C Transformed scene description data output of layer m = 1

Comprised of division candidate n = 1 and division candidate n = 2

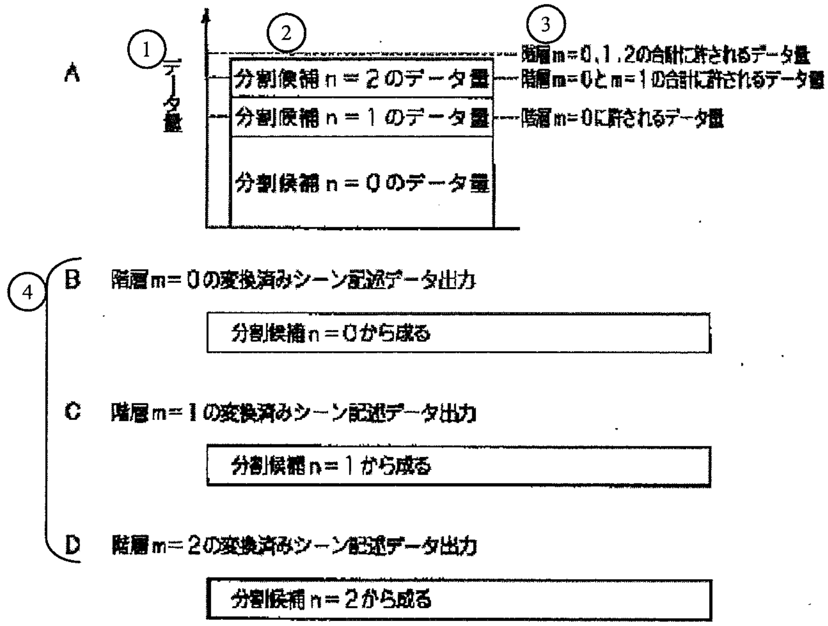


Figure 15

Key:
1 Data quantity
2 Data quantity of division candidate n = 2
Data quantity of division candidate n = 1
Data quantity of division candidate n = 0
3 Data quantity allowed for the sum of layers m = 0, 1, 2
Data quantity allowed for the sum of layers m = 0, 1
Data quantity allowed for layer m = 0
4 B Transformed scene description data output of layer m = 0
Comprised of division candidate n = 0
C Transformed scene description data output of layer m = 1
Comprised of division candidate n = 1
D Transformed scene description data output of layer m = 2
Comprised of division candidate n = 2

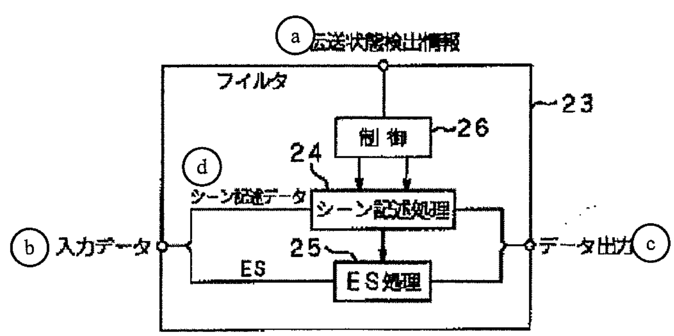


Figure 16

Key: a Transmission state detection information
b Input data
c Data output
d Scene description data
23 Filter
24 Scene description processing
25 ES processing
26 Control

	伝送優先度 (a)	ビットレート
ESa	30	Ra
ESb	20	Rb
ESc	10	Rc

Figure 17

Key: a Transmission priority
Bit rate

	伝送優先度 (a)	ビットレート
ESa	15	Ra
ESb	20	Rb
ESc	10	Rc

Figure 18

Key: a Transmission priority
Bit rate

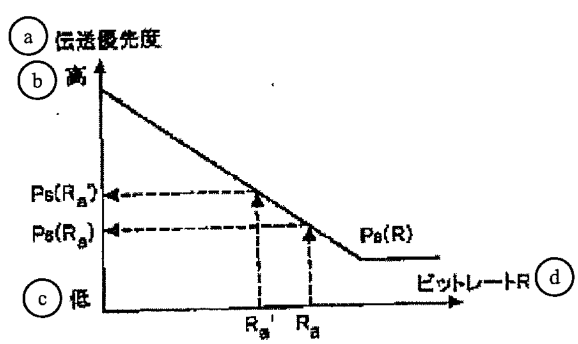


Figure 19

Key: a Transmission priority
b High

c Low
d Bit rate R

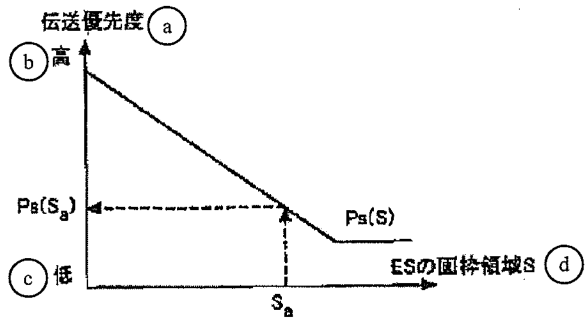


Figure 20

Key:
a Transmission priority
b High
c Low
d ES image frame region S

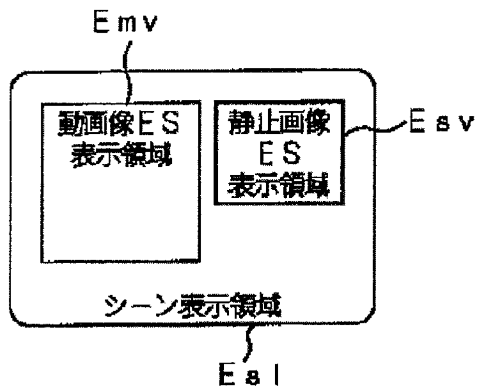


Figure 21

Key:
Emv Display region of moving picture ES
Esv Display region of still picture ES
Es1 Scene display region

```
Group{                                          a  :シーン全体のグルーピング
 children[                                         :子ノードの記述開始
  Transform{                                       :座標変換用グルーピング
   translation -1 0 0                              :local座標の原点位置#500
   scale 4 3 1                                     :local座標のscaling#501
   children[
    Shape{                                         :表示する物体
     geometry Box{}                                :立方体
     appearance Appearance{                        :物体表面特性
      texture MovieTexture{url 3}                        :テクスチャ用動画像
     }
    }
   ]
  }
  Transform{
   translation 4.5 2 0                             :local座標の原点位置#502
   scale 1 1 1                                     :local座標のscaling#503
   children[
    Shape{
     geometry Box{}
     appearance Appearance{
      texture MovieTexture{url 4}
     }
    }
   ]
  }
 ]
}
```

Figure 22

Key: a

: Grouping of overall scene
: Start of description of children node
: Grouping for coordinate transformation
: Site #500 of origin of local coordinates
: Scaling #501 of local coordinates
: Displayed object
: Cube
: Object surface characteristics
: Moving picture for texture
: Site #502 of origin of local coordinates
: Scaling #503 of local coordinates

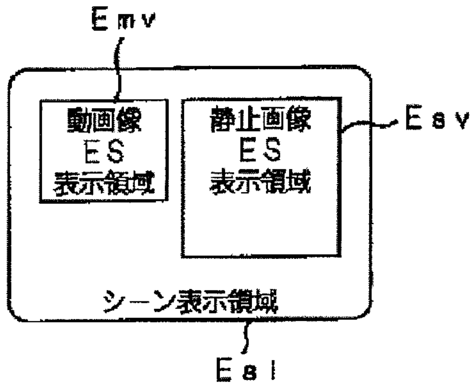


Figure 23

Key: Emv Display region of moving picture ES
Esv Display region of still picture ES
Es1 Scene display region

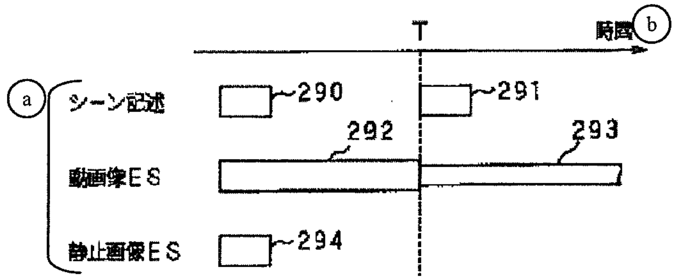


Figure 24

Key: a Scene description
Moving picture ES
Still picture ES
b Time

```
Group{                                   a  :シーン全体のグルーピング
 children[                                  :子ノードの記述開始
  Transform{                                :座標変換用グルーピング
   translation -4.5 2 0                     :local座標の原点位置#600
   scale 1 0.75 1                           :local座標のscaling#601
   children[
    Shape{                                  :表示する物体
     geometry Box{}                         :立方体
     appearance Appearance{                 :物体表面特性
      texture MovieTexture{url 3}                 :テクスチャ用動画像
     }
    }
   ]
  }
  Transform{
   translation 1 0 0                        :local座標の原点位置#602
   scale 4 4 1                              :local座標のscaling#603
   children[
    Shape{
     geometry Box{}
     appearance Appearance{
      texture MovieTexture{url 4}
     }
    }
   ]
  }
 ]
}
```

Figure 25

Key: a : Grouping of overall scene
: Start of description of children node
: Grouping for coordinate transformation
: Site #600 of origin of local coordinates
: Scaling #601 of local coordinates
: Displayed object
: Cube
: Object surface characteristics
: Moving picture for texture
: Site #602 of origin of local coordinates
: Scaling #603 of local coordinates

```
ObjectDescriptor[                        a
  ODid=3                                  :BIFS中のURL指定に対応するID
  ES_Descriptor[                          :ES情報記述子
    ES_ID=10                              :ES番号ID
    ...
    DecoderConfigDescriptor[                      :ES復号情報記述子
      ...
      bufferSizeDB=4000                           :ES復号バッファサイズ[Byte]
      maxBitRate=1000000                          :ESの最大レート[bps]
      avgBitRate=1000000                          :ESの平均レート[bps]
      ...
    }
    ...
  }
}

ObjectDescriptor[
  ODid=4
  ES_Descriptor[
    ES_ID=11
    ...
    DecoderConfigDescriptor[
      ...
      bufferSizeDB=1000
      maxBitRate=240000
      avgBitRate=240000
      ...
    }
    ...
  }
}
```

Figure 26

Key: a : ID corresponding to URL assignment in BIFS
: ES information descriptor
: ES number ID
: ES decoding information descriptor
: ES decoding buffer size (Bytes)
: ES maximum rate (bps)
: ES average rate (bps)

```
ObjectDescriptor{                    a
  ODid=3                              :BIFS中のURL指定に対応するID
  ES_Descriptor{                      :ES情報記述子
    ES_ID=10                          :ES番号ID
    …
    DecoderConfigDescriptor{                 :ES復号情報記述子
      …
      bufferSizeDB=2000                      :ES復号バッファサイズ[Byte]
      maxBitRate=500000                      :ESの最大レート[bps]
      avgBitRate=500000                      :ESの平均レート[bps]
      …
    }
    …
  }
}

ObjectDescriptor{
  ODid=4
  ES_Descriptor{
    ES_ID=11
    …
    DecoderConfigDescriptor{
      …
      bufferSizeDB=1000
      maxBitRate=240000
      avgBitRate=240000
      …
    }
    …
  }
}
```

Figure 27

Key: a : ID corresponding to URL assignment in BIFS
: ES information descriptor
: ES number ID
: ES decoding information descriptor
: ES decoding buffer size (Bytes)
: ES maximum rate (bps)
: ES average rate (bps)

```
Group[
 children[
  Transform[                          ⓐ
    translation 0 0 0              :local座標の原点位置#902
    scale 4.5 4.5 1                :local座標のscaling#903
    children[
     Shape[
      geometry Box[]
      appearance Appearance[
       texture MovieTexture[url 4]
      ]
     ]
    ]
   ]
  ]
]
```

Figure 28

Key: a : Site #902 of origin of local coordinates
: Scaling #903 of local coordinates

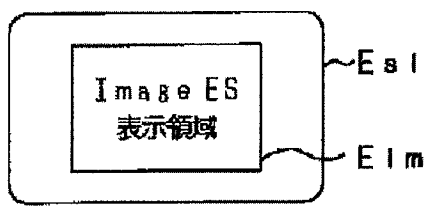


Figure 29

Key: Eim Image ES display region

```
Group{
 children[
  Transform{
   children[                          a
    Shape{                            :表示する物体
     geometry IndexedFaceSet{               :ポリゴン
      coord Coordinate{               :頂点座標のセット
       point[
        0.0, 0.0, 0.0,
        ...
        1.0, -1.0, 0.0]
      ]
      CoordIndex [                    :使用頂点セットの指定
       0, 1, 3, -1,
       ...
       109, 110, 111, -1
      ]
     }
    }
   ]
  }
 ]
}
```

Figure 30

Key: a : Displayed object
: Polygon
: Set of apex coordinates
: Assignment of apex set used

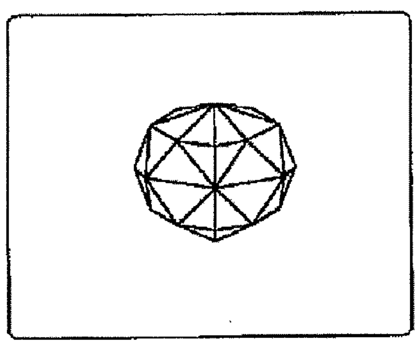


Figure 31

```
Group{
 children[
  Transform{                       a
   children[
    Shape{                        :表示する物体
     geometry Sphere{}            :球体
    }
   ]
  }
 ]
}
```

Figure 32

Key: a : Displayed object
: Ball

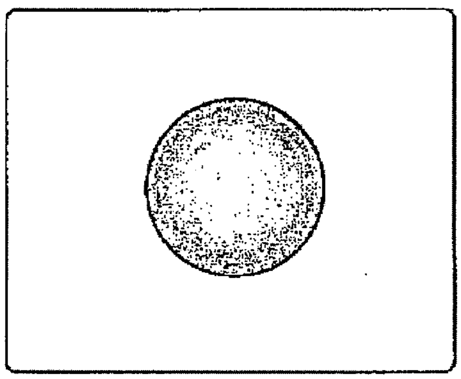


Figure 33

```
Group{ children[
   Transform{
     translation -3 0 0
     children[
       Shape{
         geometry Sphere{}          : 球体
       }
     ]
   }
  Transform{
     translation 0 0 0
     children[
       Shape{
         geometry Box{}             :立方体
       }
     ]
   }                                              a
  Transform{
     translation 0 -3 0
     children[
       Shape{
         geometry Cone{}            :円錐
       }
     ]
   }
  Transform{
     translation 3 0 0
     children[
       Shape{
         geometry Cylinder{}        :円柱
       }
     ]
   }
]}
```

Figure 34

Key: a : Ball
: Cube
: Cone
: Cylinder

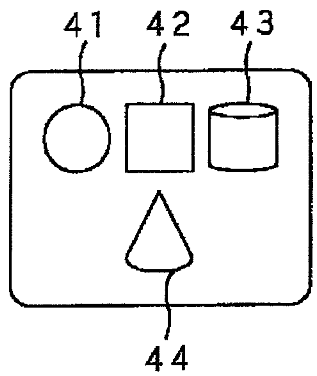


Figure 35

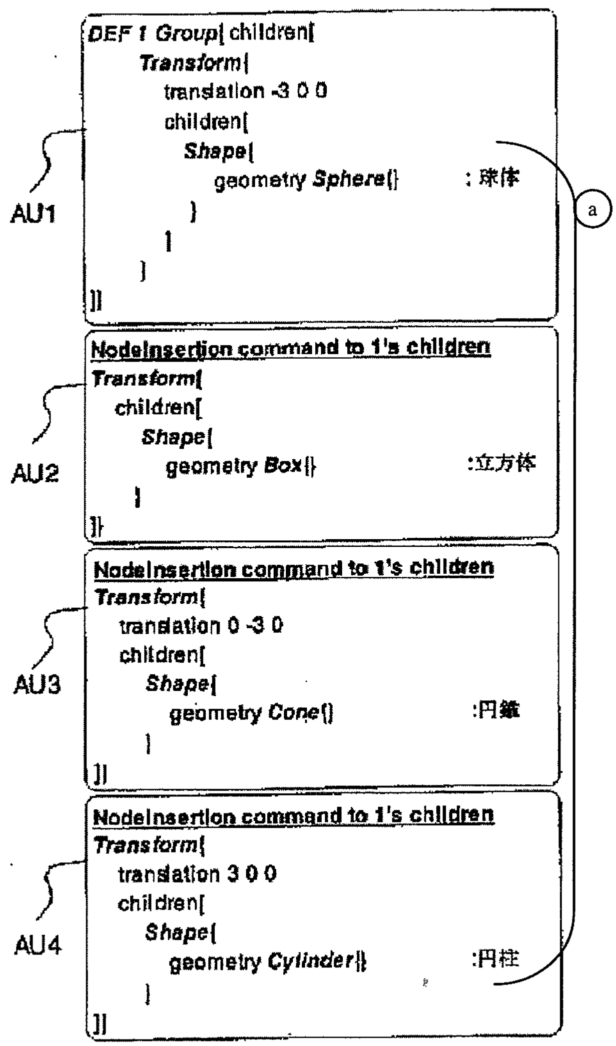


Figure 36

Key: a : Ball
: Cube
: Cone
: Cylinder

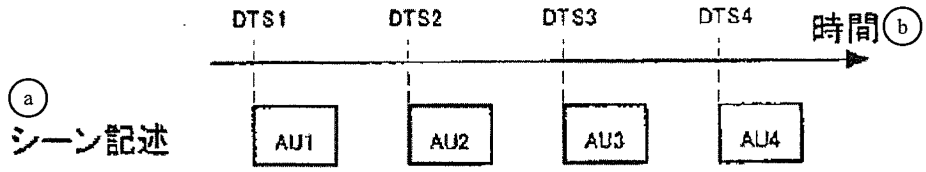


Figure 37

Key: a Scene description
b Time

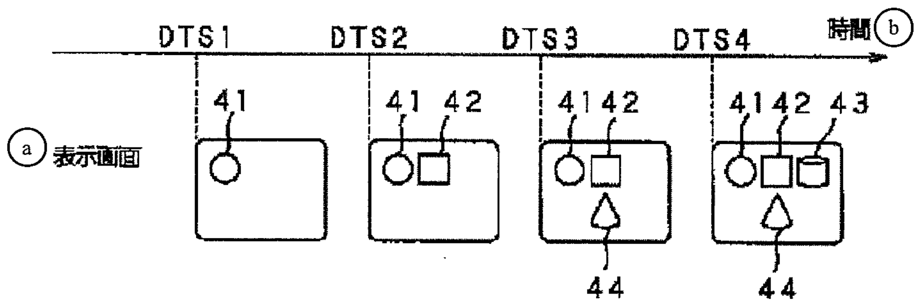


Figure 38

Key: a Displayed image
b Time

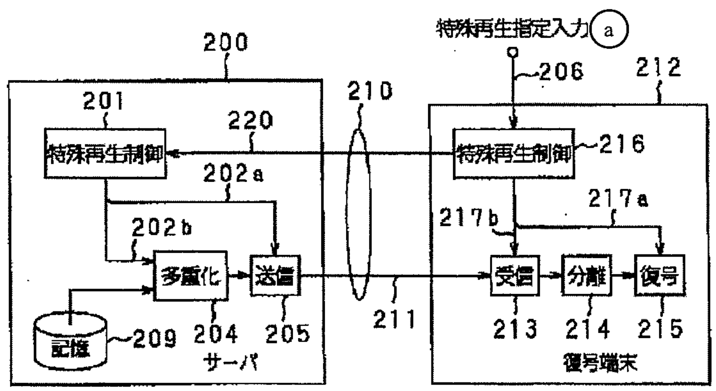


Figure 39

Key: a Input of special playback assignment
200 Server
201 Special playback control

204	Multiplexing
205	Transmission
209	Storage
212	Decoding terminal
213	Receiving
214	Separation
215	Decoding
216	Special playback control

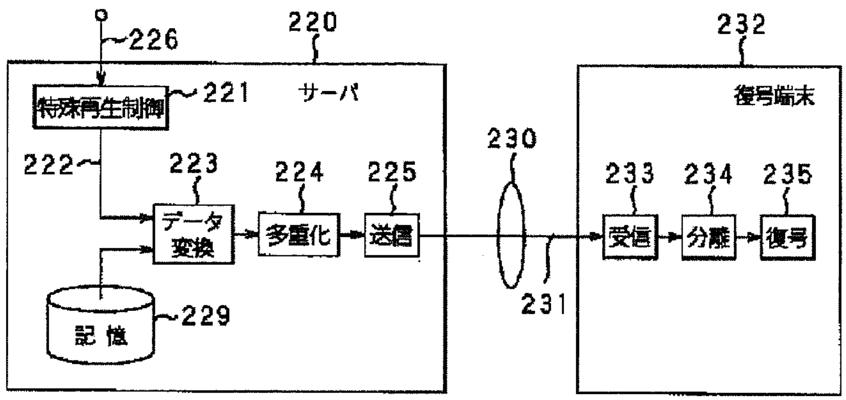


Figure 40

Key:

220	Server
221	Special playback control
223	Data transformation
224	Multiplexing
225	Transmission
229	Storage
232	Decoding terminal
233	Receiving
234	Separation
235	Decoding

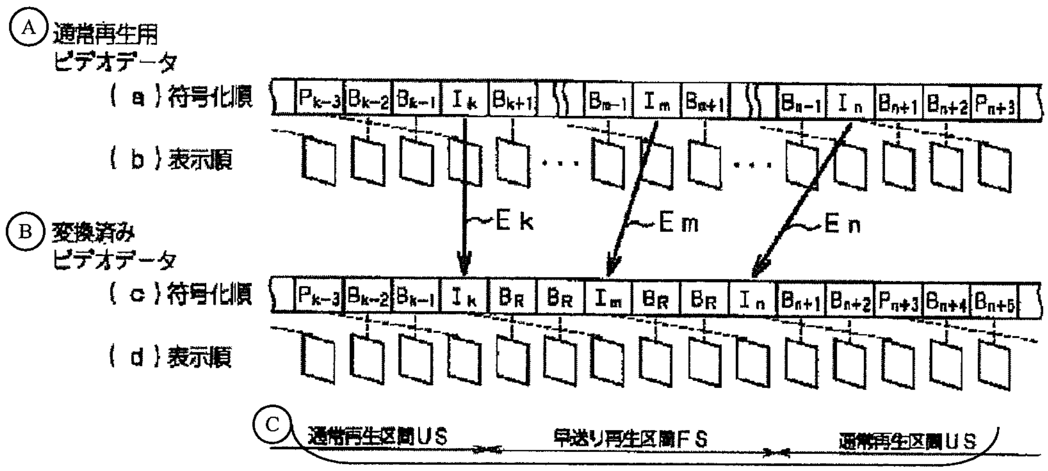


Figure 41

Key: A Video data for normal playback
(a) Coding sequence
(b) Display sequence
B Transformed video data
(c) Coding sequence
(d) Display sequence
C Normal playback interval US
FFWD play interval FS
Normal playback interval US

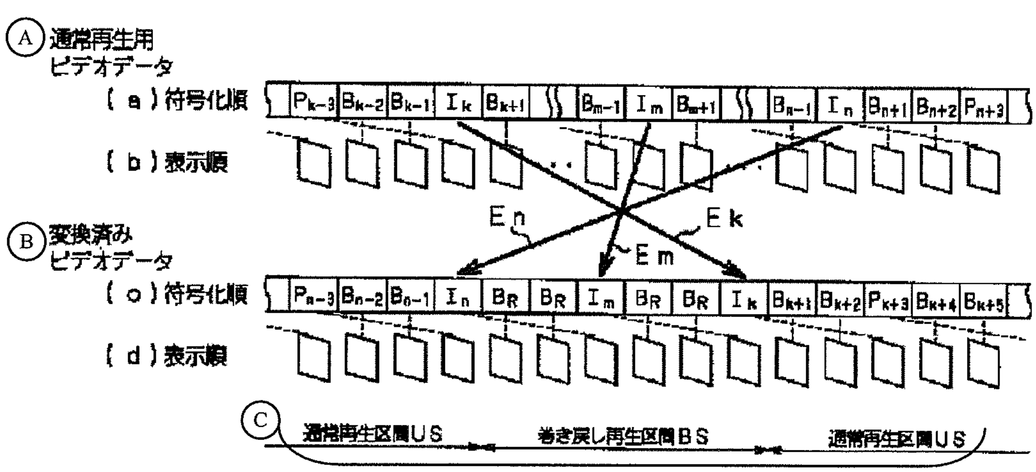


Figure 42

Key: A Video data for normal playback
(a) Coding sequence
(b) Display sequence
B Transformed video data
(c) Coding sequence
(d) Display sequence

C Normal playback interval US
REWIND play interval BS
Normal playback interval US

```
Group{                                  :シーン全体のグルーピング
 children[                              :子ノードの記述開始
  Transform{                            :座標変換用グルーピング
   translation -1 0 0                   :local座標の原点位置
   scale 4 3 1                          :local座標のscaling
   children[
    Shape{                              :表示する物体
     geometry Box{}                     :立方体
     appearance Appearance{             :物体表面特性
      texture MovieTexture{url 3}       :表面にマッピングする動画像
     }
    }
   ]
  }
  Transform{
   translation 4.5 2 0
   scale 1 1 1
   children[
    Shape{
     geometry Sphere{}                  :球体
    }
   ]
  }
 ]
}
```

Figure 43

Key: a

: Grouping of overall scene
: Start of description of children node
: Grouping for coordinate transformation
: Site of origin of local coordinates
: Scaling of local coordinates
: Displayed object
: Cube
: Object surface characteristics
: Moving picture mapped to the surface
: Ball

Continued from front page

(51) Int. Cl.^{7}:		Identification Codes:	FI		Theme Codes (for reference):
H 04 N	5/92		H 04 N	5/91	L 5C064
	7/025			5/92	H
	7/03			7/08	A
	7/035			7/13	Z
	7/24				
	7/173	610			

(72) Inventor: Yoichi Yagasaki
Sony Corp.
6-7-35 Kitashinagawa,
Shinagawa-ku, Tokyo

F-Terms (for reference):
5C025 CA02 CA09 CA18 CB10 DA01
5C025 CA02 CA09 CA18 CB10 DA01
5C052 AC03 AC05 CC06 CC11 DD10
5C053 FA23 GB37 KA04 KA05 LA11 LA15
5C059 KK37 MB02 MB22 PP01 PP04 PP19 PP20 RB02 RC19 RC32 RC33 RC34 SS02 UA02 UA05
5C063 AA01 AB03 AB07 AC02 AC05 AC10 CA20 CA23 CA29 CA31 CA36 DA02 DA03 DA05 DA07 DA13 EB01 EB37 EB39
5C064 BA01 BA07 BB05 BC10 BC16 BC23 BC25 BD08 BD09